Macro-evolutionary Studies in the Orchid Genus *Satyrium* Sw. and Other Genera from the Cape Floristic Region (South Africa)

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde

(Dr. sc. Nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Timotheüs van der Niet

aus

die Niederlände

Promotionskomitee

Prof. H. Peter Linder (Vorsitz)

Prof. Peter K. Endress

Prof. E. Conti

Zürich, 2006

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude towards my supervisor, Peter Linder. This thesis would not be what it is now without your inspiring ideas. I appreciate your style of supervision which gave me the freedom to explore the things I was interested in but, at the same time, kept me focused on the subject each time when I was about to drift away. I admire your creativity and opportunism in thinking about scientific topics and I hope I have picked up some of these features myself. Besides your professional role, you were also great company while exploring the Swiss and South African mountains. Thanks goes out to Steve Johnson who tought me important pollination-related technniques in the field and who continuously challenged me to think about his interesting hypotheses.

I would like to thank many of my colleagues who are, or were at one time, in Zurich. Particularly Chloé Galley for great company both in the field and in Zürich, Berit Gehrke who kept up with me for five weeks in Africa, Merran Matthews for correcting the English of my manuscripts over and over. And these people, plus Cyril Guibert, Julien Bachelier, Philip Moline, Frank Rutschmann, Mike Thiv, Christ Hardy, Anita Lendel, Mike Pirie, Aelys Humphries, and Marlon Machado for wonderful friendship and support in times where I needed it. It will not be easy to have to say goodbey to all of you!

Finally, I would like to thank my family and friends in the Netherlands who have continuously showed their interest in me, supported me, and visited me on a number of occassions. I would like to mention Roeland, Ina, Thomas, Sjoerd, and Mees van der Niet; Elsbeth, Johan, Tijs, Anke, and Daan Lodder; Paul, Odette, Mathilde, and Charlotte van der Niet; Niels and Leonie de Zwarte, Peter van Loon, Maurits Hofman, Iddo Schotanus, Wouter and Elske van Laar, Kees-Anton van Welie, Mark Smits and Christien de Kramer, and Wim Calame and Maria van der Barselaar. Last but not least I owe much to my parents who have supported me in many ways. I can sincerely say that without you, this project would not have been possible and I therefore would like to dedicate my work to you.

TABLE OF CONTENT

INTRODUCTION

Knowledge on the origin of biodiversity and interactions among its constituents is highly relevant in an era where such biodiversity is becoming critically threatened. Given the theory of descent with modification (Darwin 1859), one could ask for any group of species 'what drove this diversification?' In the past, addressing this question was limited by the availability of phylogenetic hypotheses (e.g. Stebbins 1950; Grant 1981). Therefore, the units that were studied were typically races or varieties within a species or, what were thought to be, closely related species. Although this approach was useful for unraveling factors that caused populations to diversify, it failed to reveal any large-scale patterns due to the absence of replication. The emergence of robust methods to infer phylogenetic relationships (Hennig 1966), to optimize attributes of organisms onto phylogenies (e.g. Swofford and Maddison 1987), and to put a temporal component onto diversification events (e.g. Sanderson 2002) has opened up the way for macro-evolutionary comparative studies (e.g. Givnish and Sytsma 1997).

The ingredients of any study of diversification include species and their attributes, which are both biotic and abiotic interactions, all of which can be classified. Taxonomy is concerned with the former, ecology with the latter. Additionally, these descriptive data can be analyzed in an evolutionary context using the above mentioned methods. Given that the result of any comparative study is influenced by the sampling of the organisms and that a common sampling strategy is one which relies on the taxonomy of a group, there is a link between the quality of the taxonomy of a group and the evolutionary interpretations that are made. Incorrect taxonomy will lead to a flawed estimation of evolutionary diversity. Classical α-taxonomy is usually based on careful investigation of morphology. It has been shown that traditional groupings of species do not necessarily reflect phylogenetic relationships due to amazing morphological convergence among species that occupy similar niches (e.g. Kay et al. 2005). And therefore these groups for which the taxonomy is largely based on suites of characters that are under convergent selection, should be treated with caution.

Therefore a first starting point for evolutionary studies should be to test whether taxonomic classifications are congruent with phylogenetic relationships. A general concern is that by using morphological characters that are potentially exposed to selective agents which act independently of phylogenetic relationships, convergence may lead to spurious morphology-based phylogenetic reconstructions (e.g. Bremer and Eriksson 1992; but see Luckow and Bruneau 1997).For this reason, DNA sequence data from neutral markers are most commonly used to reconstruct phylogenetic relationships (e.g. Givnish and Sytsma

1997). Additionaly, this type of data allows for putting a temporal component to diversification events as well (e.g. Sanderson 2002).

Comparative studies that aim to test for associations between organismal attributes should be carried out in a phylogenetic framework to controls for phylogenetic inertia (Felsenstein 1985). Therefore reconstruction of a species tree is a necessary ingredient for any study of diversification (Barraclough and Nee 2001). However, phylogenetic trees that result from analysis of DNA sequences from single genes cannot be equated to a species tree (e.g. Doyle 1992, 1997; Maddison 1997). This is because the phylogenetic history of a particular gene may not be congruent with the phylogenetic history of that particular organism, or because of analytical artefacts that are caused by specific properties of the dataset. Therefore reconstruction of a species tree should involve thorough sampling of DNA sequence data from several genomes and careful examination of cladograms derived from separate phylogenetic analyses.

Initial observations of the ecology and morphology can result in hypotheses as to what may have driven diversification of a taxon (e.g. Schluter 1996). A combination of a species tree and the relevant attributes of the species, allows for the testing of associations of certain attributes with diversification events (e.g. Goldblatt and Manning 1996), and their influence on diversification rates (Sanderson and Donoghue 1996).

In the first three chapters of this thesis I present the results of a macro-evolutionary study where I address these three main issues on congruence between taxonomy and phylogenetic relationships, reconstruction of a species tree, and what factors may have driven diversification. I have addressed the issues using the terrestrial orchid genus *Satyrium*. Its 90 species are distributed throughout southern, eastern and western Africa, and Madagascar, with four species extending into Asia (Summerhayes 1968a; Summerhayes 1968b; Bose and Bhattacharjee 1980; Geerinck 1984; Polunin and Stainton 1984; la Croix and Cribb 1995; Cribb and Thomas 1997; Wood 1997; Chen et al. 1999; Linder and Kurzweil 1999). *Satyrium* harbours a great deal of morphological diversity in both vegetative as well as floral characters (Kurzweil and Linder 1998). Furthermore, the species occupy a wide range of habitats such as (montane) grassland, coastal dune thicket, fynbos, and even forest (Kurzweil and Linder 1998). Finally, a wide range of pollination systems have been described (Johnson 1997; Johnson and Steiner 1999) including all the syndromes *sensu* Fægri and Van der Pijl (1979). These attributes make *Satyrium* an attractive system in which to study the driving forces for diversification.

The infrageneric classification of *Satyrium* has been problematic, with each taxonomic revision different from the preceding one (Lindley 1830-1840; Bolus 1889; Kränzlin 1901; Schlechter 1902; Kurzweil and Linder 1998). In chapter 1 I address the following question using DNA sequence data: are the subgenera of *Satyrium* monophyletic?

The aim of chapter 2 is to reconstruct a species tree for *Satyrium.* I use DNA sequence data from the plastid and nuclear genome. Phylogenetic incongruence among separate gene trees compromises straight forward reconstruction of the species tree. Given the observation of extensive phylogenetic incongruence, I specifically address the following questions: (1) are any of the observed cases of incongruence significant? (2) Is the incongruence the result of non-biological artifacts such as insufficient taxon sampling or long-branch attraction, or (3) can we assign incongruence to biological phenomena such as orthology/paralogy conflation, lineage sorting and hybridization? Finally, (4) what is the best species tree for *Satyrium*?

In chapter 3 I explore the interaction between morphology and pollinator shifts in a phylogenetic context. Pollinator observations exist for some 25 species (e.g. Johnson 1997). We combined these observations with a large dataset of floral characters that are putatively involved in pollination and habitat data to address the following questions: (1) what is the relationship between floral characters and taxa pollinated by the same pollinator class, and the phylogeny respectively? (2) Which, if any, floral characters evolve in a correlated fashion with shifts to certain pollinator classes? (3) How can we best infer pollinator classes for the taxa for which pollinator observations are lacking? (4) How often do shifts to different pollinator classes occur and are shifts to a certain pollinator class biased from other pollinator classes? (5) Are diversification rates among the different pollinator classes the same? (6) Is there a bias among taxa pollinated by a certain pollinator class for to occur in certain habitats and are pollinator shifts associated with habitat shifts?

By investigating large-scale diversification patterns within an evolutionary lineage, it is possible to detect replications of similar evolutionary events. This approach can alternatively be applied to geographical areas instead (Pennington et al. 2004) . The question then becomes: are there any features intrinsic to an area that have driven diversification? This replication can be detected by comparing patterns of diversification for several lineages within the area. In chapter 4, I describe such a study for the Cape Floristic Region (CFR) in southwestern South Africa. The CFR is characterized by a large species diversity and high levels of endemism (Linder 2003). Many lineages show a large diversity in floral form. It is believed that this is the result of adaptation to different pollinators (e.g. Johnson and Steiner 1997; Goldblatt and Manning 2000). What still remains to be tested is the precise role of a

shift in pollination system in the speciation process and, more specifically, whether (1) it reflects adaptation to different pollinators according to Stebbin's principle of the 'most effective pollinator' (Stebbins 1970; Johnson 1996) or (2) whether it has come about to acquire reproductive isolation, by selection against hybrids after an initial diversification on different, but adjacent, soils (e.g. Goldblatt and Manning 1996). In chapter 4, I addressed this issue by testing for an association between pollination system shifts and edaphic shifts for sister species with overlapping distribution ranges and allopatric sister species in the CFR respectively. Furthermore I explicitly test for an association between a joint shift in both pollination system and edaphic conditions and overlapping distribution ranges.

LITERATURE CITED

Barraclough, T. G., and S. Nee. 2002. Phylogenetics and speciation. Trends Ecol. Evol. 16: 391-399.

Bolus H. 1889. Contribution to South African Bot.. Bot. J. Linn. Soc. 25: 156-210.

Bose T. K., and S. K. Bhattacharjee. 1980. *Orchids of India.* Naya Prokash, Calcutta.

Bremer, B., and O. Eriksson. 1992. Evolution of fruit characters and dispersal modes in the tropical family Rubiaceae. Biol. J. Linn. Soc. 47: 79-95.

Chen S., Z. Tsi, and Y. Luo. 1999. *Native orchids of China in colour*. Science Press, Beijing.

Cribb P. J., and S. Thomas. 1997. Orchidaceae. In: *Flora of Ethiopia and Eritrea Vol. 6* eds. Edwards, S., S. Demissew, and I. Hedberg. Addis Ababa and Uppsala.

Darwin, C. 1859. *On the origin of species by means of natural selection.* London: John Murray.

Doyle, J. J. 1992. Gene trees and species trees – molecular systematics as one-character taxonomy. Syst. Bot. 17: 144-163.

Doyle, J. J. 1997. Trees within trees: genes and species, molecules and morphology. Syst. Biol. 46: 537-553.

Fægri, K., and L. Van der Pijl. 1979. *The principles of pollination ecology.* Oxford: Pergamon Press.

Felsenstein, J. 1985. Phylogenies and the comparative method. Am. Nat. 125: 1-15.

Geerinck D. 1984. Orchidaceae (première partie). In: *Flore d'Afrique Centrale* ed. Bamps, P.. Meise: Jardin botanique national de Belgique.

Givnish, T. J., and K. J. Sytsma (eds.). 1997. *Molecular evolution and adaptive radiation.* New York: Cambridge University Press.

Goldblatt, P. and J. C. Manning. 1996. Phylogeny and speciation in *Lapeirousia* subgenus *Lapeirousia* (Iridaceae: Ixioideae). Ann. Mo. Bot. Gard. 83: 346-361.

Goldblatt, P. and J. C. Manning. 2000. The long-proboscid fly pollination system in southern Africa. Ann. Mo. Bot. Gard. 87: 146-170.

Grant, V. 1981. *Plant speciation.* New York: Columbia University Press.

Hennig, W. 1966. *Phylogenetic systematics.* Urbana: University of Illinois Press.

Johnson, S. D. 1996. Pollination, adaptation and speciation models in the Cape flora of South Africa. Taxon 45: 59-66.

Johnson, S.D. 1997. Insect pollination and floral mechanisms in South African species of *Satyrium* (Orchidaceae). Plant Syst. Evol. 204: 195-206.

Johnson, S. D. and K. E. Steiner. 1997. Long-tongued fly pollination and evolution of floral spur length in the *Disa draconis* complex (Orchidaceae). Evolution 51: 45-53.

Johnson, S. D., and K. Steiner. 1999. Pollination. Pp. 34-39 in: *Orchids of Southern Africa* eds. Linder, H. P., and H. Kurzweil. Rotterdam: Balkema.

Kay, K. M., P. A. Reeves, R. G. Olmstead, and D. W. Schemske. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus* (Costaceae): evidence from nrDNA ITS and ETS sequences. Am. J. Bot. 92: 1899-1910.

Kränzlin F. 1901. *Orchidearum genera et species*. Mayer and Müller, Berlin.

Kurzweil, H., and H. P. Linder. 1998. A phylogenetic analysis of the genus *Satyrium* (Orchidaceae). *Beiträge zur Biologie der Pflanzen* 71: 101-181.

La Croix I., and P. J. Cribb. 1995. Orchidaceae. In: *Flora Zambesica Vol. 11 part 1* ed. Pope, G. V. London : Flora Zambesica Managing Committee.

Linder, H. P. 2003. The radiation of the cape flora, southern Africa. Biol. Rev. 78: 597-638.

Linder, H. P., and H. Kurzweil. 1999. *Orchids of Southern Africa*. Rotterdam: Balkema.

Luckow, M., and A. Bruneau. 1992. Circularity and independence in phylogenetic tests of ecological hypotheses. Cladistics 13: 145-151.

Lindley J. 1830-1840. *The genera and species of Orchidaceaous plants*. London: Ridgways.

Maddison, W.P. 1997. Gene trees in species trees. Syst. Biol. 46: 523-536.

Pennington, R. T., Q. C. B. Cronk, and J. A. Richardson. 2004. Introduction and synthesis: plant phylogeny and the origin of major biomes: one contribution of 16 to a Discussion Meeting Issue ‘Plant phylogeny and the origin of major biomes’. Philos. Trans. R. Soc. Lond., B, Biol. Sci. 359: 1455-1464.

Polunin O., and A. Stainton. 1984. *Flowers of the Himalaya*. Bombay: Oxford University Press.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. 19; 101-109.

Sanderson, M. J., and M. J. Donoghue. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol. Evol. 11: 15-20.

Schlechter R. 1902. Monographie der Diseae. Bot. Jahrb. Syst. Pflanzengesch. Pflanzengeogr. 31: 134-313.

Schluter, D. 1996. Ecological causes of adaptive radiation. Am. Nat. 148: S40-S64.

Stebbins jr., G. L. 1950. *Variation and evolution in plants.* New York: Columbia University Press.

Stebbins Jr., G. L. 1970. Adaptive radiation of reproductive characteristics in angiosperms, I: pollination mechanisms. Annu. Rev. Ecol. Syst. 1: 307-326.

Summerhayes V. S. 1968a. Orchidaceae (part 1). In: *Flora of Tropical East Africa* eds. Milne-Redhead, E., and R. M. Polhill. London: Crown Agents for Oversea Governments and Adminstrations.

Summerhayes V. S. 1968b. Orchidaceae Vol III Part I. In: *Flora of West Tropical Africa* eds. Hutchinson, J., J. M. Dalziel and F. N. Hepper. London: Crown Agents for Oversea Governments and Adminstrations.

Swofford, D. L., and W. P. Maddison. 1987. Reconstructing ancestral character states under Wagner parsimony. Math. Biosci 87: 199-229.

Wood J. R. I. 1997. *A handbood of the Yemen Flora*. Kew: Royal Botanic Gardens.

Molecular Markers Reject Monophyly of the Subgenera of *Satyrium* (Orchidaceae)

TIMOTHEÜS VAN DER NIET,[1,3] H. PETER LINDER,[1] BENNY BYTEBIER[2], AND DIRK U. BELLSTEDT[2]

[1]*Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, CH-8008, Zurich, Switzerland;*

[2]*University of Stellenbosch, Biochemistry Department, Private Bag X1, 7602 Matieland, South Africa;*

Systematic Botany 2005, 30:263-274

Abstract.—*Satyrium* is a morphologically anomalous genus. Although clearly a member of the terrestrial orchid subfamily Orchidoideae, its phylogenetic relationships within the subfamily are uncertain. Morphologically it has always been placed in tribe Diseae, albeit associated with different subtribes, but recent molecular studies suggest a closer relationship to Orchideae. The infrageneric classification of *Satyrium* is equally problematic, and several different classifications are available. The only infrageneric classification based on a phylogenetic analysis of morphological characters resulted in recognition of the three subgenera *Brachysaccium, Bifidum* and *Satyrium*. DNA sequence data from nuclear (ITS1, 5.8S, and ITS2) and plastid (*trnL* intron, *trnL-F* intergenic spacer, and part of the *matK* gene and *trnK* intron) genome are used to test the monophyly of these subgenera. Topologies of cladograms resulting from parsimony analysis of separate datasets show several cases of incongruence, some of which are well supported. Combined analysis is performed on a dataset from which two problematic taxa are pruned. Parametric bootstrap, as well as Bayesian posterior probability, rejects monophyly of all three subgenera and alternative groupings are suggested.

Keywords: incongruence, infrageneric classification, ITS, parametric bootstrap, plastid, Satyrium

INTRODUCTION

Satyrium Sw. is a terrestrial orchid genus consisting of 89 species distributed primarily throughout southern, eastern and western Africa and Madagascar with four species extending into Asia (Fig. 1) (Humbert 1939; Summerhayes 1968a, 1968b; Bose and Bhattacharjee 1980; Geerinck 1984; Polunin and Stainton 1984; la Croix and Cribb 1995; Cribb and Thomas 1997; Wood 1997; Chen et al. 1999; Linder and Kurzweil 1999). Together with *Pachites* Lindl., a genus with two rare species endemic to the southwestern tip of South Africa, it constitutes the subtribe Satyriinae Schltr. (Schlechter 1926; Dressler 1993; Linder and Kurzweil 1994; Kurzweil and Linder 2001). Both genera possess non-resupinate flowers, simple petals and sepals and a gynostemium with a long basal column-part and a pendent anther. *Satyrium* is characterized by a galeate labellum with two spurs, and *Pachites* has a subactinomorphic perianth (Kurzweil and Linder 2001). Satyriinae was included in Diseae Dressler although precise affinities within the tribe remained obscure (Linder 1986; Linder and Kurzweil 1994; Kurzweil et al. 1995). Molecular studies indicate a close relationship between Satyriinae and Orchideae (Dressler and Dodson) P. Vermeulen, thereby rendering Diseae paraphyletic (Cameron et al. 1999; Douzery et al. 1999; but see Bellstedt et al. 2000). There is no unequivocal evidence for placement of Satyriinae as sister to any of the tribes or subtribes within Orchideae. However, monophyly of the subtribe seems well established, although most studies supporting this did not include either of the two members of *Pachites*.

Satyrium harbors a great deal of morphological diversity, especially in plant size, leaf shape and orientation, flower color, spur length, and gynostemium features (Summerhayes 1968a; Hall 1982; Kurzweil 1996; Kurzweil and Linder 1998). Obtaining a stable infrageneric classification has consequently proved difficult. Every taxonomist who has studied the group proposed a more or less different scheme (Table 1) (Lindley 1830-1840; Bolus 1889; Kränzlin 1901; Schlechter 1902; Kurzweil and Linder 1998). Lindley (1830-1840) even recognized the monotypic *Aviceps* Lindl. and *Satyridium* Lindl. in addition to *Satyrium* based on their autapomorphic characters. Cladistic analysis of morphological characters by Kurzweil and Linder (1998) resulted in the most recent infrageneric classification, recognizing three subgenera: *Brachysaccium* Kurzweil and Linder, *Bifidum* Kurzweil and Linder and *Satyrium*. Support for the classification is poor, although the subgenera can be diagnozed morphologically (Table 2). Kurzweil and Linder (1998) noted '…that one of the groups cannot be monophyletic on the current data' and that their study 'highlights the need for more data sets to resolve the patterns within *Satyrium*.'

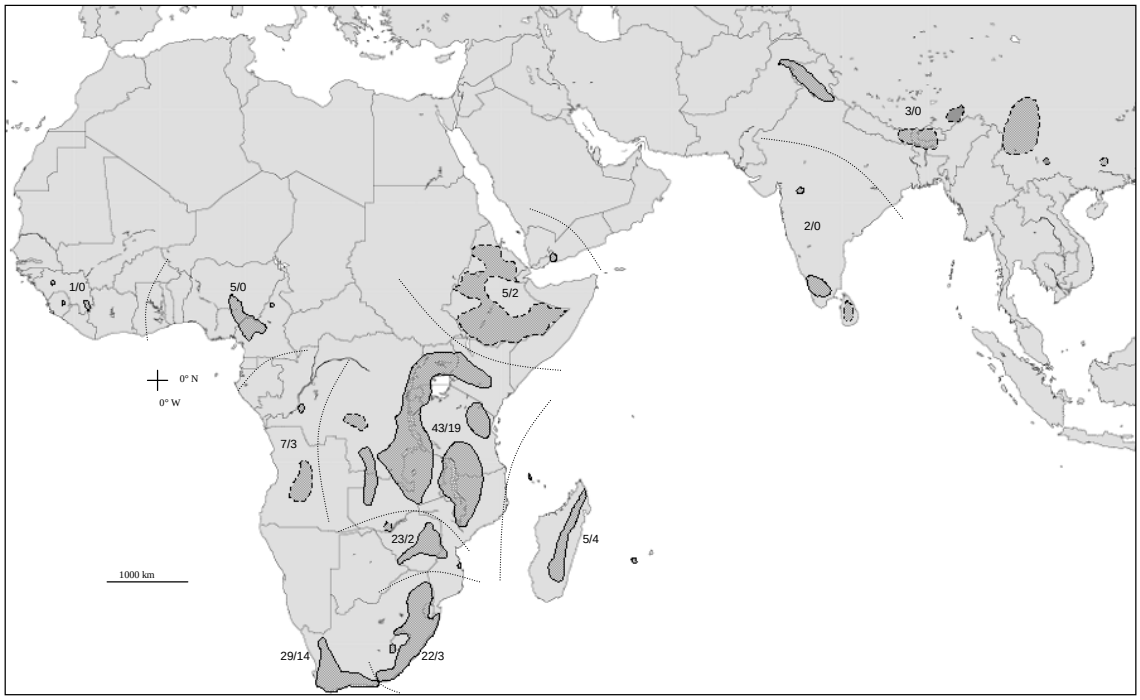


FIG. 1. Distribution map of *Satyrium*. Areas for which accurate distribution data are available are indicated within continuous lines. If distribution data are inaccurate, areas are indicated by dashed lines. For each region between dotted lines the number of species (left from slash) and the number of endemics (right from slash) respectively is given.

This study tests the monophyly of the subgenera of *Satyrium* by using molecular markers and addresses the debate over a robust infrageneric classification. As interspecific hybridization has been described in *Satyrium* (Hall 1982; Ellis and Johnson 1999), DNA sequence data from both nuclear (ITS1, 5.8S, and ITS2) and plastid (*trnL* intron and *trnL-F* intergenic spacer and part of the *matK* gene and *trnK* intron) genome were used.

MATERIAL AND METHODS

Taxon Sampling

Thirty-eight species of *Satyrium* with representatives from all previously recognized groups, including *Aviceps* and *Satyridium*, were sampled throughout the distribution range of the genus (Table 3). This includes the Cape Floristic Region (CFR, sensu Goldblatt 1978), the Drakensberg, East Africa, Madagascar, and Southeast Asia. Species-level taxonomy follows

TABLE 1. Taxonomic treatment by various taxonomists of taxa included in this study. Names follow genera (printed in bold) if taxon was placed outside *Satyrium* and sections (normal font) of *Satyrium* of Lindley (1830-1840), subgenera (printed in bold) if taxon was placed outside the mains subgenus *Satyrium* and sections (normal font) of the subgenus *Eusatyrium* of Bolus (1889), sections (printed in bold) if taxon was placed outside section *Eusatyrium* and subsections (normal font) of section *Eusatyrium* of Kränzlin (1901), sections of Schlechter (1902) and subgenera of Kurzweil and Linder (1998). Brackets indicate that the species was listed as a synonym. If two names are listed, species and a conspecific taxon were placed into two different sections.

Species	Lindley (1830-1840)	Bolus (1889)	Kränzlin (1901)	Schlechter (1902)	Kurzweil and Linder (1998)
Satyrium humile	*Longicalcarata*	*Humistratae*	*Bifolia*	*Eu-Satyrium*	*Bifidum*
S. cristatum		*Adscendentes*	*Coriifolia*	*Leptocentrum*	*Bifidum*
S. erectum	*Brevicalcarata*	*Humistratae*	*Macrophylla*	*Leptocentrum*	*Bifidum*
S. amoenum	*Longicalcarata*		*(Bifolia), Macrophylla*	*Leptocentrum*	*Bifidum*
S. orbiculare			*Bifolia*		*Bifidum*
S. pallens					*Bifidum*
S. princeae					*Bifidum*
S. breve			*Trinervia*	*(Leucocomus)*	*Brachysaccium*
S. pumilum	***Aviceps***	*Saccatae*	***Aviceps***	*Aviceps*	*Brachysaccium*
S. bicallosum	*Saccata*	*Saccatae*	*Bracteata*	*Brachysaccium*	*Brachysaccium*
S. microrrhynchum			*Bifolia*	*Brachysaccium*	*Brachysaccium*
S. bracteatum	*Saccata*	*Saccatae*	*(Bifolia), Bracteata*	*Brachysaccium, (Eu-Satyrium)*	*Brachysaccium*
S. muticum	*Saccata*	*Saccatae*	*Bifolia*	*Eu-Satyrium*	*Brachysaccium*
S. trinerve	*Saccata*	*Adscendentes*	*(Coriophoroidea), Trinervia*	*Leucocomus*	*Brachysaccium*
S. amblyosaccos					*Brachysaccium*
S. shirense				*(Chlorocorys)*	*Satyrium*
S. volkensii			*(Coriophoroidea)*	*(Chlorocorys)*	*Satyrium*
S. ligulatum	*Brevicalcarata*	*Adscendentes, (Humistratae)*	*(Bracteata), Coriifolia*	*(Eu-Satyrium), Leptocentrum*	*Satyrium*
S. crassicaule			*Macrophylla*	*(Imperfectly known), Leptocentrum*	*Satyrium*

(Table 1 continued)

S. parviflorum	*(Brevicalcarata), Longicalcarata*	*Adscendentes*	*Coriophoroidea*	*Chlorocorys*	*Satyrium*
S. odorum		*Adscendentes*	*Coriifolia*	*Chlorocorys*	*Satyrium*
S. bicorne	*(Longicalcarata)*	*Humistratae*	*Bifolia*	*Eu-Satyrium*	*Satyrium*
S. acuminatum	*Longicalcarata*	*Humistratae*	*Bifolia*	*Eu-Satyrium*	*Satyrium*
S. carneum	*Longicalcarata*	*Humistratae*	*Bifolia*	*Eu-Satyrium*	*Satyrium*
S. membranaceum	*Longicalcarata*	*Humistratae*	*Bifolia*	*Eu-Satyrium*	*Satyrium*
S. coriifolium	*Brevicalcarata*	*Adscendentes*	*Coriifolia*	*Leptocentrum*	*Satyrium*
S. longicauda	*Longicalcarata*	*Adscendentes*	*Macrophylla*	*Leptocentrum*	*Satyrium*
S. stenopetalum	*Longicalcarata*	*Adscendentes*	*Cofiifolia*	*Leptocentrum*	*Satyrium*
S. hallackii		*Adscendentes*	*Macrophylla*	*Leptocentrum*	*Satyrium*
S. nepalense	*(Brevicalcarata), Longicalcarata*		*Coriifolia*	*Leptocentrum*	*Satyrium*
S. ciliatum	*Brevicalcarata*		*(Coriifolia)*	*Leptocentrum*	*Satyrium*
S. buchananii			*(Imperfectly known)*	*Leptocentrum*	*Satyrium*
S. rupestre			*Coriophoroidea*	*Leptocentrum*	*Satyrium*
S. rhynchanthum	***Satyridium***	***Satyridium***	***Satyridium***	*Satyridium*	*Satyrium*
S. chlorocorys			*Coriophoroidea*		*Satyrium*
S. microcorys					*Satyrium*
S. sceptrum					*Satyrium*
S. sphaeranthum					*Satyrium*

TABLE 2. Distribution of some morphological characters accross the three subgenera of *Satyrium* showing the wealth of diversity. If two states are indicated, both are present in the subgenus.

Character	Brachysaccium	Bifidum	Satyrium
sterile shoots	present/absent	present/absent	present/absent
leaf position	cauline/basal	cauline/basal	cauline/basal
inflorescence	lax/dense	lax/dense	lax/dense
fusion sepals and petals	quarter to half of length	weak up to half of length	weak to extensive
lip galea	mostly wide entrance	wide entrance	wide or narrow entrance
spur type	saccate	slender	slender
spur length	mostly shorter than ovary	mostly longer than ovary	mostly longer than ovary
rostellum	lateral lobes spreading or parallel	lateral lobes parallel	weakly trilobed or unlobed
viscidia	terminal	terminal	lateral

TABLE 3. List of taxa included in this study with country of origin, voucher information and Genbank accession numbers. Herbarium abbreviations follow Holmgren and Keuken (1974). Abbreviations are as follows: BB = Benny Bytebier; HK = Hubert Kurzweil; TvdN = Timo van der Niet, CH = Switzerland, MA = Malawi, MD = Madagascar, SA = South Africa, TA = Tanzania.

Dactylorhiza maculata (L.) Soo, TvdN218, CH (Z): (*trnL-F* AY705005 and AY705045, *matK* AY708007, ITS AY704973). ***Gymnadenia conopsea*** (L.) R.Br., TvdN219, CH (Z): (*trnL-F* AY705006 and AY705046, *matK* AY708008, ITS AY704974). ***Platanthera chlorantha*** Cust. ex Reichb., TvdN222, CH (Z): (*trnL-F* AY705007 and AY705047, *matK* AY708009, ITS AY704975). ***Satyrium acuminatum*** Lindl., TvdN18b, SA (Z): (*trnL-F* AY705008 and AY705048, *matK* AY708010), HK1854, SA (NBG): (ITS AJ000142). ***Satyrium amblyosaccos*** Schltr., HK2004, MA (MAL, UZL): (*trnL-F* AY705049 and AY705009, *matK* AY708010, ITS AY704976). ***Satyrium amoenum*** A.Rich., Hermans 5401, MD (K): (*trnL-F* AY705010 and AY705050, *matK* AY708012, ITS AY704977). ***Satyrium bicallosum*** Thunb., BB2112, SA (BR, NBG): (*trnL-F* AY705011 and AY705051), HK1883, SA (NBG): (*matK* AY708013, ITS AJ000143). ***Satyrium bicorne*** (L.) Thunb., TvdN46, SA (Z): (*trnL-F* AY705012 and AY705052, *matK* AY708014, ITS AY704978). ***Satyrium bracteatum*** (L.f.) Thunb., BB2110, SA (BR, K, NBG, NY): (*trnL-F* AY705013 and AY705053, *matK* AY708015, ITS AY704979), BB2191, SA (BR, GRA, NBG): (*trnL-F* AY705014 and AY705054, *matK* AY708016, ITS AY704980). ***Satyrium breve*** Rolfe, BB2175, TA (EA): (*trnL-F* AY705015 and AY705055, *matK* AY708017, ITS AY704981). ***Satyrium buchananii*** Schltr., HK2053, MA (MAL): (*trnL-F* AY705016 and AY705056, *matK* AY708018, ITS AY704982). ***Satyrium carneum*** (Dryand.) Sims, TvdN3, SA (Z): (*trnL-F* AY705017 and AY705057, *matK* AY708019), HK1814, SA (NBG): (ITS

(Table 3 continued)

AJ000136). ***Satyrium chlorocorys*** Reich.f. ex Rolfe, HK1969, MA (MAL, PRE, SRGH, UZL): (*trnL-F* AY705018 and AY705058, *matK* AY708020, ITS AY704983). ***Satyrium ciliatum*** Lindl., Luo and Luo 728 (unknown): (*trnL-F* AY714744 and AY714748, *matK* AY714746, ITS AY714743). ***Satyrium coriifolium*** Sw., TvdN47 (Z): (*trnL-F* AY705019 and AY705059, *matK* AY708021, ITS AY704984). ***Satyrium crassicaule*** Rendle, HK2030, MA (MAL): (*trnL-F* AY705020 and AY705060, *matK* AY708022, ITS AY704985). ***Satyrium cristatum*** Sond., BB2297, SA (GRA, NBG): (*trnL-F* AY705021 and AY705061, *matK* AY708023, ITS AY704986). ***Satyrium erectum*** Sw., BB2062, SA (BR, K, NBG, NY, Z): (*trnL-F* AY705022 and AY7050652, *matK* AY708024, ITS AY704987). ***Satyrium hallackii*** Bolus, BB2258, SA (BR, K, NBG): (*trnL-F* AY705023 and AY705063, *matK* AY708025, ITS AY704988). ***Satyrium humile*** Lindl., TdvN22, SA (Z) (*trnL-F* AY705024 and AY705064, *matK* AY708026), HK1884, SA (NBG): (ITS AJ000134). ***Satyrium ligulatum*** Lindl, HK1815, SA (NBG): (*trnL-F* AY705025 and AY705065 , ITS AJ000141), TvdN6, SA (Z): (*matK* AY708027). ***Satyrium longicauda*** Lindl., BB2249, SA (BR, NBG): (*trnL-F* AY705026 and AY705066, *matK* AY708028, ITS AY704989). ***Satyrium membranaceum*** Sw., HK1822, SA (NBG): (*trnL-F* AY705027 and 705067, *matK* AY708029), HK1834, SA (NBG): (ITS AJ000144). ***Satyrium microcorys*** Schltr., HK2015, MA (LMA, MAL, PRE, SRGH, UZL): (*trnL-F* AY705028 and 705068, *matK* AY708030, ITS AY704990). ***Satyrium microrrhynchum*** Schltr, BB2276, SA (NBG): (*trnL-F* AY705029 and AY705069, *matK* AY708031, ITS AY704991). ***Satyrium muticum*** Lindl., WL802-1 (NBG): (*trnL-F* AY705030 and AY705070, *matK* AY708032, ITS AY704992). ***Satyrium nepalense*** D. Don, Chase O-539, unknown (K): (*trnL-F* AY714745 and AY714749, *matK* AY714747, ITS AJ000140). ***Satyrium odorum*** Sond., TvdN59, SA (Z): (*trnL-F* AY705031 and AY705071, *matK* AY708033), HK1811, SA (NBG): (ITS AJ000133). ***Satyrium orbiculare*** Rolfe, HK2043, MA (MAL): (*trnL-F* AY705032 and AY705072, *matK* AY708034, ITS AY704993). ***Satyrium pallens*** S.D.Johnson & H.Kurzweil, TvdN21, SA (Z): (*trnL-F* AY705033 and AY705073, *matK* AY708035, ITS AY704994). ***Satyrium parviflorum*** Sw., BB2217, SA (BR, GRA, NBG): (*trnL-F* AY705034 and AY705074, *matK* AY708036, ITS AY704995). ***Satyrium princeae*** Kraenzl., HK2005, MA (MAL, SRGH, UZL): (*trnL-F* AY705035 and AY705075, *matK* AY708037, ITS AY704996). ***Satyrium pumilum*** Thunb., BB2012, SA (BR, NBG, Z): (*trnL-F* AY705036 and AY705076, *matK* AY708038, ITS

AY704997). ***Satyrium rhynchanthum*** Bolus, K Steiner s.n., SA (NBG): (*trnL-F* AY705037 and AY705077, ITS AJ000130), BB2155, SA (BR, NBG, Z): (*matK* AY708039). ***Satyrium***

(Table 3 continued)

rupestre Schltr., DUB503, SA (NBG): (*trnL-F* AY705038 and AY705078, *matK* AY708040, ITS AY704998). ***Satyrium sceptrum*** Schltr., HK1985, MA (MAL, UZL): (*trnL-F* AY705039 and AY705079, *matK* AY708041, ITS AY704999). ***Satyrium shirense*** Rolfe, HK1968, MA (LMA, MAL, PRE, SRGH, UZL): (*trnL-F* AY705040 and AY705080, *matK* AY708042, ITS AY705000). ***Satyrium sphaeranthum*** Schltr., HK1990, MA (MAL, UZL): (*trnL-F* AY705041 and AY705081, *matK* AY708043, ITS AY705001). ***Satyrium stenopetalum*** Lindl., BB2096, SA (BR, NBG, Z): (*trnL-F* AY705042 and AY705082, *matK* AY708044, ITS AY705002). ***Satyrium trinerve*** Lindl., BB2255, SA (BR, NBG): (*trnL-F* AY705043 and AY705083, *matK* AY708045, ITS AY705003). ***Satyrium volkensii*** Schlechter, BB2177, TA (BR, EA): (*trnL-F* AY705044 and AY705084, *matK* AY708046, ITS AY705004).

Kurzweil and Linder (1998). Two accessions of *Satyrium bracteatum* (L.f.) Thunb. that represent morphologically and geographically diverse forms were included. ITS sequences of nine species were downloaded from Genbank (Douzery et al. 1999). Sequences from plastid loci of some of these species were subsequently obtained from different accessions, but these species were treated as composite terminals in the combined analysis. Three European members of Orchideae were used as outgroup to root the cladograms. In the absence of material of *Pachites*, phylogenetic position of its two species remains unaddressed here.

Molecular Techniques and Data Matrix Composition

DNA was extracted from silica-dried or fresh leaf material using either Dneasy Plant Mini Kits (Quiagen, Basel, Switzerland) following the manufacturer's protocol or the CTAB method (Doyle and Doyle 1987). Polymerase chain reactions (PCR) were performed in a total volume of 25 µl in a Biometra Thermocycler or TGradient (Biometra, Göttingen, Germany) using 2.5 mM $MgCl_2$, 1x PCR buffer (Amersham Biosciences), 0.25 mM of dNTPs and 0.1 mM (*trnL* intron and *trnL-F* intergenic spacer), 0.32 mM (part of the *matK* gene and *trnK* intron) and 0.4 mM (ITS) of each primer and 1 unit of *Taq* DNA polymerase (Amersham Biosciences). The *trnL* intron and *trnL-F* intergenic spacer (hereafter simply *trnL-F*) was amplified using the "f" primer (Taberlet et al. 1991) and primer C3 (5'-GCAGAGACTCAATGGAAGCTG) annealing approximately 150 base pairs (bp) downstream of the "c" primer (Taberlet et al. 1991) to avoid sequencing of a single nucleotide

repeat string (Bellstedt et al. 2001). PCR procedure included 35 cycles of 1 min denaturation at 94ºC, 1 min annealing at 50ºC, and 2 min extension at 72ºC. Part of the *matK* gene and *trnK* intron (hereafter simply *matK*) was amplified using one of the forward primers F19, 580F or 1082F and R1 (Kocyan et al. in press). PCR procedure included an initial 3 min of denaturation at 95ºC followed by 34 cycles of 30 sec denaturation at 95ºC, 1 min annealing at 52ºC, 1 min 40 sec extension at 72ºC, and was terminated by a 7 min final extension at 72ºC. ITS1, 5.8S, and ITS2 (hereafter simply ITS) was amplified using primers ITS5 and ITS4 (White et al. 1990). PCR procedure included addition of 1 µl of DMSO (Sigma-Aldrich, St. Louis, MO, USA) and 25 cycles of 1 min denaturation at 94ºC, 1 min annealing at 54ºC, and 2 min 30 sec extension at 72ºC with two additional seconds elongation per cycle. PCR products were purified using either the QIAquick PCR Purification Kit (Qiagen, Basel, Switzerland) or GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences). Cycle sequencing reactions were carried out with the ABI Prism Big Dye Terminator Cycle Sequencing Ready Extraction Kit (Applied Biosystems, Foster City, CA, USA). Sequencing reactions were purified on MicroSpin G-50 columns (Amersham Biosciences) and loaded on an ABI PRISM 3100 genetic analyzer (Applied Biosystems, Foster City, CA, USA). Sequences were assembled and edited using Sequencher™ Version 3.1.1 (GeneCodes Corporation). Double peaks in the electropherograms were coded as polymorphism according to the IUPAC coding. Most of the alignment was done by eye. Highly variable parts of the sequence flanked by conserved parts were aligned using CLUSTAL W version 1.8 (Thompson et al. 1994) and adjusted manually. Gaps were coded using the 'simple indel coding' method (Simmons and Ochoterena 2000) as implemented in Gapcoder (Young and Healy 2003). In the final matrix 0.04% of the cells was coded as 'missing data'. As the plastid genome is inherited as a unit, incongruence among sequences due to different phylogenetic histories is improbable and therefore the *trnL-F* and *matK* sequences were readily combined into a single plastid dataset. A copy of the complete dataset can be obtained from TreeBASE (study accession number S1134; matrix accession number M1945).

Parsimony Analysis

Nuclear (ITS) and plastid (*trnL-F* and *matK*) datasets were first analyzed separately. All characters were equally weighted and treated as unordered. Heuristic searches using PAUP***** 4.0b (Swofford 2000) were started with 1000 stepwise random addition sequence replicates, holding 10 trees at each step, followed by TBR swapping, saving maximally 100 most parsimonious trees per replicate with MULTREES and steepest descent in effect**.** All

shortest cladograms retained in memory were input for a second round of searching involving exhaustive TBR swapping. Character support for each node was inferred using the bootstrap procedure (Felsenstein 1985). Five hundred bootstrap replicates were performed with 250 stepwise random addition sequence replicates, 3 trees held each step, saving no more than 10 trees, with MULTREES and steepest descent in effect.

To maximize the explaining power of the available data, the ITS and plastid dataset were combined (Kluge 1989; Nixon and Carpenter 1996). However, these data partitions potentially reflect different phylogenetic histories. In a combined phylogenetic analysis this incongruence violates the assumption of a bifurcating tree (Bull et al. 1993; de Queiroz 1993; de Queiroz et al. 1995; Hillis 1995; Miyamoto and Fitch 1995). Therefore, first congruence among topologies derived from sequence data from ITS and plastid genome was assessed using the Partition Homogeneity Test (PHT) as implemented in PAUP* 4.0b (Swofford 2000). Each PHT involved 100 replicates and each individual replicate involved the same heuristic search strategy as applied to the bootstrap. The combined dataset was partitioned into genome-specific partitions. Subsequently, that set of taxa causing incongruence based on visual inspection of topologies derived from analysis of genome-specific partitions, were pruned from the combined dataset. The hypothesis that the pruned datamatrix contained no incongruent taxa was tested by the PHT. The pruned taxa were then individually added to the combined dataset to detect if their inclusion would cause rejection of the null hypothesis of congruence. Taxa that caused rejection of the null hypothesis of congruence were removed from all analyses involving combination of the ITS and plastid dataset. For analysis of the combined dataset, tree search and assessment of support involved similar routines as described for the separate datasets.

Bayesian Inference

Bayesian inference was performed using MrBayes v. 3.0 (Ronquist and Huelsenbeck 2003). Data were partitioned into functional categories (ITS1, 5.8S, ITS2, *trnL* intron, *trnL-F* intergenic spacer, *matK*, and gaps). The most optimal model of sequence evolution for each of these partitions was selected using Modeltest version 3.06 (Posada and Crandall 1998). A combined dataset of all partitions, including gaps, was analyzed applying separate models to each data partition and with parameters estimated separately for each individual partition. One million generations were run with parameters sampled every 1,000 generations. Based on inspection of the likelihood scores for each generation, the first 250 generations were

considered as burnin. This analysis involved the same taxon sampling as used for the parsimony analysis of the combined dataset.

Hypothesis Testing

The hypothesis that the larger number of steps required for a cladogram containing the subgenera defined by Kurzweil and Linder (1998), compared to the most parsimonious cladogram, was simply due to stochasticy of the process of sequence evolution was tested using the parametric bootstrap (Hillis et al. 1996; Huelsenbeck et al. 1996). Null distributions for the parametric bootstrap analyses were generated by constraining the subgenera to be monophyletic. Sites resulting from gaps, missing data, and ambiguity were excluded. Modeltest version 3.06 (Posada and Crandall 1998) was run for each subgenus separately to select the optimal model of sequence evolution and to estimate the parameter values of this model. These were used to estimate branch lengths of the constrained topology and to simulate 100 datasets onto the constrained topology with branch lengths using Seq-Gen.v1.2.6 (Rambaut and Grassly 1997). The datasets were subject to parsimony analyses with and without the topological constraint of monophyly of the individual subgenera enforced. The difference in number of steps between these two scores was plotted in a frequency diagram. The observed value from parsimony analysis of the original data matrix excluding gaps, missing data, and ambiguous sites was then compared to the obtained frequency distribution, and significance ($P<0.001$) was assessed. The posterior probability of monophyletic subgenera was read from the output file of MrBayes v 3.0 with posterior probability values of clades (Ronquist and Huelsenbeck 2003).

RESULTS

The ITS dataset yielded a high number and percentage of parsimony informative characters. Of the 693 positions in the aligned matrix, 296 were parsimony informative (Table 4). The ITS dataset was further characterized by 78 gaps, ranging in size from 1-28 bp (Table 4). Two different *matK* sequences were obtained for *Satyrium crassicaule* Rendle, depending on the PCR method. Both sequences were aligned in separate matrices and subject to parsimony analysis. If analyzed together with the *trnL-F* sequences, similar topologies were obtained. The sequence that provided the shortest tree was selected for further analysis. Lengths of the 57 inferred gaps for the plastid dataset ranged from 1-182 bp and were typically largest in *trnL-F*. Twenty gaps in *trnL-F*, resulting from ambiguous alignment of

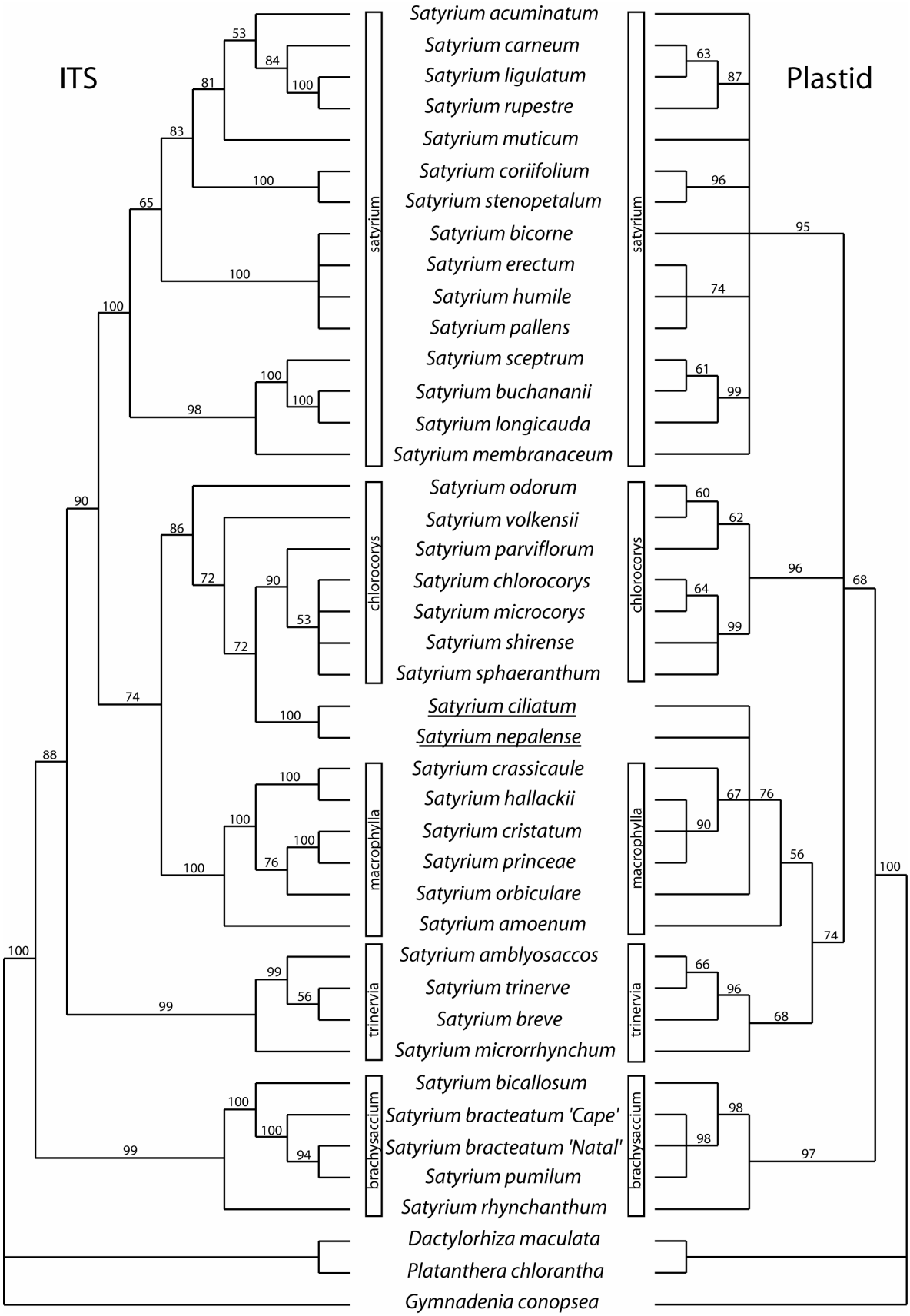


FIG. 2. Strict consensus cladograms derived from separate analysis of ITS and plastid dataset. Bootstrap values are highlighted above the branches. The two taxa that were excluded from combined analyses due to conflicting position between ITS and plastid topology are underlined. The two accessions of *Satyrium bracteatum* are labeled ‘Cape’ and ‘Natal’ representing their geographic origin. Informal clade names of present study are indicated in bars next to respective clades.

TABLE 4. Sequence statistics of all datasets used in parsimony analysis.

Data partition	ITS	*matK*	*trnL-F*	Plastid	Combined
Sequence length	617-650	628-663	610-791	-	-
Aligned length	693	702	998	-	-
Number of parsimony informative characters (nucleotides only)	296	49	39	-	-
% of parsimony informative characters (nucleotides only) calculated over all nucleotides	42,7	7	3,9	-	-
Number of gaps	78	11	46	-	-
% of parsimony informative gaps calculated over all gaps	54	36	42	-	-
% of nodes supported by >70% BS support	77	25	25	44	66

repeat regions, were excluded from all analyses. The percentage of parsimony informative characters was lower for the plastid dataset than for the ITS dataset (Table 4).

Parsimony analysis of the ITS dataset resulted in 48 most parsimonious cladograms (Fig. 2) with length (L) = 860 steps, consistency index (CI) = 0.57, and retention index (RI) = 0.80 (all statistics excluding autapomorphies). Analysis of the plastid dataset retrieved 528 most parsimonious cladograms (Fig. 2) with L = 161 steps, CI = 0.71, and RI = 0.89. The percentage of branches supported by at least 70% bootstrap support (BS), was much higher for the ITS dataset than for the plastid dataset (Table 4). Comparison of the topologies derived from analysis of ITS and plastid dataset revealed several cases of incongruence. Most of these received only low bootstrap support (low bootstrap support was de fined here as 50%-85% BS, high bootstrap support was defined here as 86%-100% BS) from at least one of the datasets. The PHT indicated that the conflicting position of the Asian *Satyrium ciliatum* Lindl. and *Satyrium nepalense* D. Don was significant (Table 5). As a result of the observed incongruence, these taxa were removed from all combined analyses.

TABLE 5. Taxa placed in conflicting position in the cladograms based on ITS and plastid dataset with the P-values of the PHT resulting from adding these taxa individually to a combined data matrix from which conflicting taxa have been pruned. Taxa marked with an asterisk were excluded from all combined analyses based on this P-value.

Taxon	P-value
Satyrium amblyosaccos, Satyrium breve, Satyrium microrrhynchum, Satyrium trinerve	0.32
Satyrium buchananii	0.96
Satyrium carneum	0.90
Satyrium crassicaule	0.88
Satyrium ciliatum, Satyrium nepalense**	0.01
Satyrium odorum	0.84
Satyrium orbiculare	0.32
Satyrium parviflorum	0.98
No conflicting taxon added	1.00

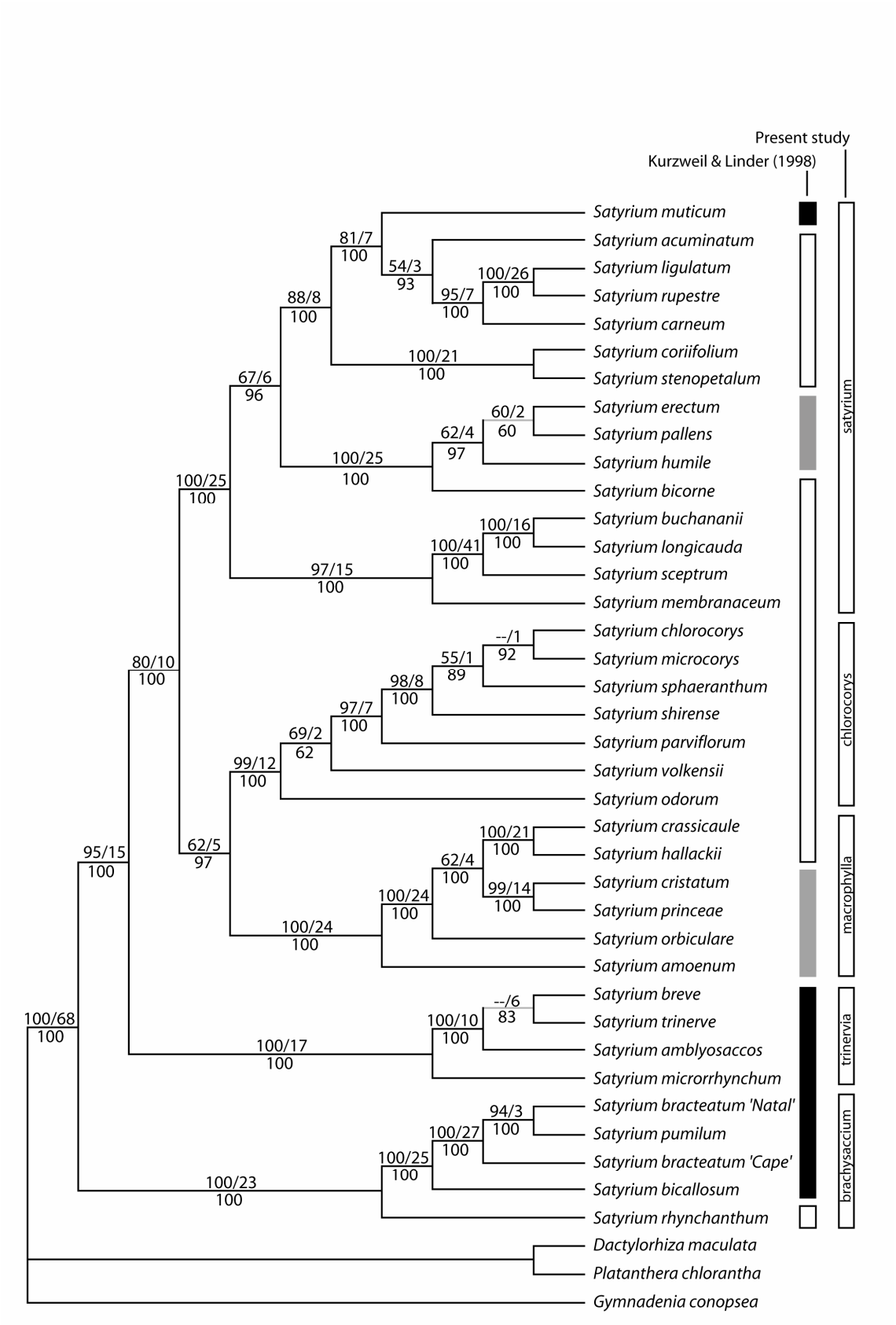


FIG. 3. Phylogenetic tree from Bayesian analysis with posterior probability below branches and parsimony bootstrap and branch lengths (DELTRAN) values above branches. Two nodes that collapse in the strict consensus cladogram of parsimony analysis are highlighted in light gray. Subgenera of Kurzweil & Linder (1998) are highlighted in black bars (*Brachysaccium*), gray bars (*Bifidum*) and white bars (*Satyrium*). Informal clade names of present study are indicated in bars next to respective clades.

Parsimony analysis of the combined dataset excluding *S. ciliatum* and *S. nepalense* resulted in an almost fully resolved strict consensus cladogram (Fig. 3) with L = 1000 steps, CI = 0.60 and RI = 0.81. Most nodes were well supported.

For the Bayesian analysis the the model GTR + γ was selected for *matK, trnL-F* intergenic spacer, ITS1, 5.8S, and ITS2 whereas F81 + γ for the *trnL* intron was selected by Modeltest. Bayesian analysis resulted in a similar topology although posterior probability values for clades were higher than parsimony bootstrap values (Fig. 3). In the remainder of the paper the strict consensus cladogram resulting from parsimony analysis of the combined dataset and its BS values (Fig. 3) will be used for discussion unless mentioned otherwise.

Monophyly of *Satyrium* was supported by 100% BS (Fig. 3). Within *Satyrium*, five main clades (here labeled with informal names), most of which receive high bootstrap support from both separate and combined analysis, were recognized. The brachysaccium clade containing species mostly occurring in the CFR was sister to all other species. Monophyly of the brachysaccium clade, as well as that of its sister clade, was supported by both datasets though the latter clade only received 68% BS in the plastid dataset (Fig. 2). Both *Aviceps* (*Satyrium pumilum* Thunb.) and *Satyridium* (*Satyrium rhynchanthum* Bolus) were included in the brachysaccium clade. *S. bracteatum* was paraphyletic with respect to *S. pumilum*. The highly supported (100% BS) trinervia clade containing species with relatively short and saccate spurs, branched off next. This position was congruent with the ITS topology (Fig. 2). In the plastid topology this clade was sister to the macrophylla clade (Fig. 2). The PHT test could not reject the null hypothesis of congruence between the ITS and plastid datasets despite the observation that the grouping of the trinervia clade and macrophylla clade as sisters was supported independently by both plastid loci. A big clade, containing three well supported subclades (macrophylla, chlorocorys, and satyrium clade respectively) was supported as sister by 80% BS (Fig. 3). Species with slender spurs from Madagascar and tropical Africa (the macrophylla clade) grouped together with the chlorocorys clade, but this relationship was supported by only 62% BS (Fig. 3). Not all nodes in the chlorocorys clade received high bootstrap support. Within the macrophylla clade, apart from the sister group relationship of *S. crassicaule* + *Satyrium hallackii* Bolus and *Satyrium princeae* Kraenzl. + *Satyrium cristatum* Sond, relationships were all well supported, including the position of the Madagascan *Satyrium amoenum* A. Rich. as sister to the rest (Fig. 3). The fifth main clade (the satyrium clade) was retrieved in all analyses. Within this clade, the basal split represented a poorly supported sister group relationship between species from mainly tropical Africa and the CFR. Not all of the internal nodes were well supported.

Parametric bootstrap (Fig. 4) rejected for all three subgenera the null hypothesis that the observed difference between constraint cladograms containing monophyletic subgenera and the most parsimonious cladograms was caused by stochasticy of the substitution process (P<<0.001). Posterior probability values of monophyletic subgenera were significantly small (P<<0.001).

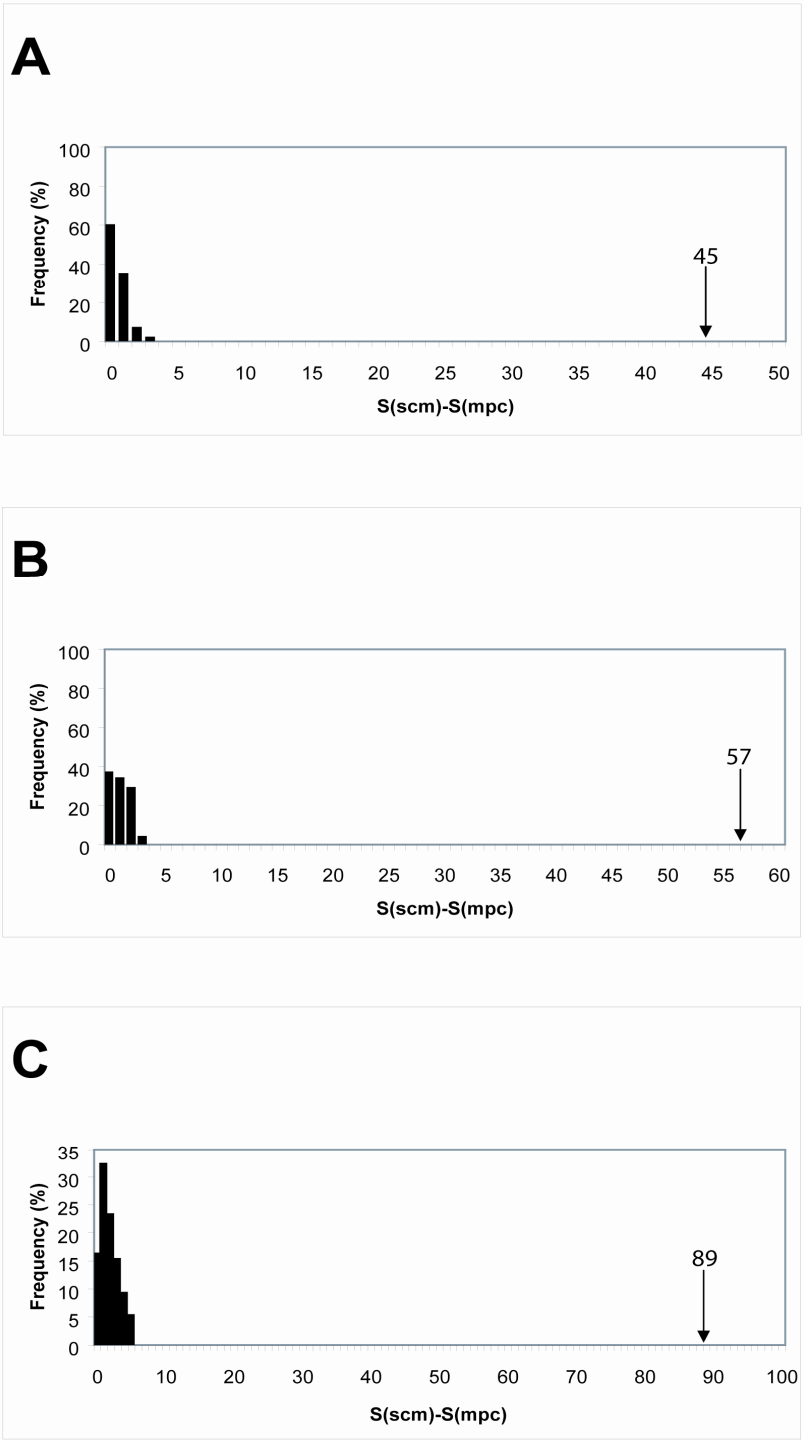


FIG. 4. Frequency diagrams with the null distribution of $Steps_{(subgenerus\ constrained\ monophyletic)}$ - $Steps_{(most\ parsimonious\ cladogram)}$ (S(scm)-S(mpc)) generated by parametric bootstrap and the observed value highlighted with an arrow for: A. *Brachysaccium*. B. *Bifidum*. C. *Satyrium*.

DISCUSSION

There is a marked difference in number of parsimony informative characters for the loci used in this study. The nuclear ITS sequences provide both a larger absolute number and a higher proportion of parsimony informative characters compared to the plastid loci (Table 4). This is also reflected in the higher percentage of nodes supported by at least 70% bootstrap support. This pattern is found in several studies of Orchidaceae using DNA sequences (Whitten et al. 2000; Gravendeel et al. 2001; Pridgeon et al. 2001; Salazar et al. 2003), while it has been observed only occasionally in other land plants (Lavin et al. 2001; Moody et al. 2001; Van der Heede et al. 2003). A possible artifact of the seemingly higher number of parsimony informative characters provided by ITS here could be that both *matK*, and *trnL-F* are characterized by large gaps. These inflate the number of characters over which the percentage of parsimony informative characters is calculated. However, recalculating the statistics on our matrices with gaps excluded, returns a similar result (the biggest change is in *trnL-F*, from 3.9% to 5.6%). An explanation for this could be that most of the nucleotide variation rests in regions that also exhibit extensive length variation.

There is topological incongruence between the ITS and plastid datasets. Apart from human error, incongruence is associated with processes at the molecular level (Wendel and Doyle 1998). It can be divided into two categories that have different consequences for inclusion of taxa in combined analyses. Topological conflict that stems from stochastic processes of sequence evolution ("stochastic incongruence") should not preclude the combination of datasets including all taxa, whereas if different branching histories underlie incongruence ("real incongruence"), taxa causing conflict should be excluded prior to phylogenetic analyses that rely on methods assuming bifurcation (e.g. Bull et al. 1993; de Queiroz 1993; de Queiroz et al. 1995; Hillis 1995; Miyamoto and Fitch 1995).

To discriminate between these two kinds of incongruence, many studies rely on visual inspection of the topologies and comparison of bootstrap support of incongruent nodes (Eldenäs and Linder 2000; Whitten et al. 2000; Gravendeel et al. 2001; Pridgeon et al. 2001). More explicit methods apply statistical tests to identify "real incongruence" but the most commonly used method, the PHT test implemented in PAUP* 4.0b (Swofford, 2000) or the ILD test (Farris et al. 1994), has been critized (Reeves et al. 2001; Yoder et al. 2001; Goldblatt et al. 2002). The problem is that the null hypothesis of congruence is often rejected, even if the rival topologies show only few cases of weakly supported incongruence (Reeves et al. 2001; Goldblatt et al. 2002). Such "stochastic incongruence" should not prevent combination of datasets even though the null hypothesis of congruence is rejected by the

PHT. One explanation is that “stochastic incongruence” in different parts of a cladogram may have a cumulative effect resulting in rejection of global congruence which was shown in our study. If taxa were added to the pruned congruent matrix individually, the PHT would not reject the null hypothesis of congruence, whereas if added simultaneously a significant result would be obtained. Therefore, in order to detect local incongruence the alternative application of the PHT as proposed in this paper should be used. Here, it resulted in exclusion of only *S. ciliatum* and *S. nepalense*, the Asian representatives of *Satyrium*. Their phylogenetic relationships remain ambiguous.

Monophyly of the subgenera of *Satyrium*, as delimited by Kurzweil and Linder (1998), is rejected by the molecular data (Fig. 4), using the parametric bootstrap test. As Huelsenbeck et al. (1996) regard this test as conservative, the rejection of monophyletic subgenera is convincing. The molecular data also falsify all other previous classifications (Fig. 3; Table 1) that were often based on few morphological key characters, thus indicating the need for a new infrageneric classification for the genus.

Satyridium (*S. rhynchanthum*) is included in *Satyrium*. It was treated as a monotypic genus by Lindley (1830-1840) and Hall (1982) whereas Bolus (1889), Kränzlin (1901) and Schlechter (1902) recognized it as a monotypic section of *Satyrium*. Kurzweil and Linder (1998) included it in subgenus *Satyrium* (Table 1). The gynostemium is unusual compared to that of the rest of the genus. It protrudes from the lip galea, pointing forwards. The stigma is a papillose pad situated near the top of the gynostemium below the rostellum. The gynostemium has a single viscidium (Lindley 1830-1840; Kurzweil 1996). More light might be shed on the evolution of the peculiar morphology by the inclusion of species of *Pachites* as they share some features (Kurzweil 1996).

Aviceps (*S. pumilum*) (Lindley 1830-1840; Kränzlin 1901; Bolus 1889; Schlechter 1901) was separated from *Satyrium* as it was thought to lack sepals and petals (Lindley 1830-1840), an interpretation which Kurzweil (1996) refuted. The sepals and petals are fused in front of the galea where they form a landing platform for the pollinator. Although *S. pumilum* is endemic to the CFR, in the phylogeny it is sister to the Natal specimen of *S. bracteatum*, and not to the specimen from the CFR. This is unexpected from a biogeographical point of view. The taxonomy of *S. bracteatum* is confusing as shown by the many synonyms. Hall (1982) recognized the variation in vegetative and floral features but found it too continuous to merit subdivision. More extensive sampling of this species for studies using molecular techniques might shed light on its taxonomic circumscription. Besides morphological diversity of the brachysaccium clade, Hall (1982) also mentioned the strongly deviating

chromosome numbers (n=21 for most South African *Satyrium* species, n=18 for *Satyrium bicallosum* Thunb., 2n=74 for *S. pumilum*, and n=54 for *S. bracteatum*). He suggested that this could be a clade composed of ancient relicts.

Members of the trinervia clade have been associated with members of the brachysaccium clade by Lindley (1830-1840) and Kurzweil and Linder (1998) based on their saccate spurs and rostellum structure. However, Kränzlin (1901) and Schlechter (1902) already put them in a separate sections (Table 1) acknowledging their typical morphological characters such as compact inflorescences, long spreading bracts (Kränzlin 1901), and linear-lanceolate leaves (Schlechter 1902). Even though the sister relationship is not clear from the molecular data due to incongruence, their monophyly is strongly supported. They present an interesting biogeographic pattern with the endemic *Satyrium microrrhynchum* Schltr. from the Eastern Cape in South Africa as sister to a clade of East African species including the widespread *Satyrium trinerve* Lindl.that also occurs in Southern Africa, West Africa, and on Madagascar.

The strongly supported chlorocorys clade consists mostly of species included in Kränzlin's subsection *Coriophoroidea* (1901) and Schlechter's section *Chlorocorys* (1902).

The association of species in the macrophylla clade, occurring in tropical Africa, the summer rainfall region of South Africa and Madagascar was not suggested before, although Schlechter (1902) put *S. amoenum*, *S. cristatum*, *S. hallackii*, and *S. crassicaule* in a section together with several other species based on leaf position and the absence of a contracted galea entrance. Molecular data support a sister relationship between *S. crassicaule* and *S. hallackii*. They occupy wet habitats and share other features such as pink flowers and elongate leaves arising from the base (Summerhayes 1968a; Hall 1982). *S. cristatum* and *S. princeae* represent a sister species pair with a disjunct distribution and a different leaf shape (upright leaves vs. leaves flat appressed to the ground) and flower color (white, tinged red vs. pink). Sampling of additional species might identify other species as closer relatives.

Some of the relationships within the satyrium clade are remarkable from a morphological point of view. The small, white-flowered *Satyrium stenopetalum* Lindl. is in a well-supported sister-relationship to the robust orange-flowered *Satyrium coriifolium* Sw. The same applies to the large flesh colored flowers of *Satyrium carneum* (Dryand.) Sims which is sister to *Satyrium ligulatum* Lindl. and *Satyrium rupestre* Schltr. ex Bolus with minute, whitish flowers. A more thorough study of morphology may identify characters that could explain these relationships.

The convincing falsification of previous classifications puts the use of few morphological key characters to group taxa into question. In the light of the molecular data, traditionally used characters are often shown to be uninformative for diagnosing clades. Bolus (1889) and Schlechter (1902) used leaf arrangement to delimit groups (Bolus 1889; Schlechter 1902), but Kurzweil and Linder (1998) regarded this character as being too variable (even within species, e.g. *Satyrium longicauda* Lindl.). However, if optimized onto the cladogram, it can be shown that the evolution of a pair of basal leaves spreading on the ground supports the branch subtending the satyrium clade (Fig. 5).

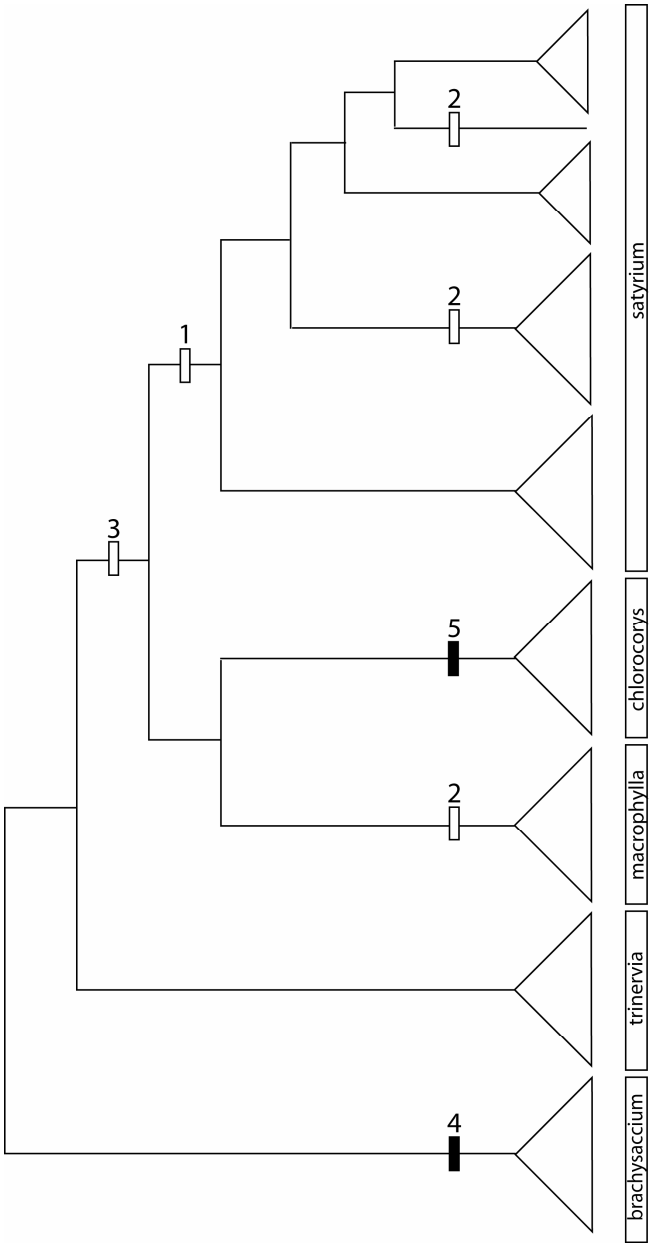


FIG. 5. Some key morphological characters optimized onto the strict consensus cladogram of combined analysis of nuclear and plastid data. Closed bars indicate non-homoplasious transitions, open bars indicate homoplasious transitions. 1. Foliage leaves: cauline/basal. 2. Rostellum: weakly trilobed or unlobed/bifid. 3. Spur type: saccate/slender. 4. Anticlinal wall of seed coat: straight/undulate. 5. Galea entrance: wide/contracted. All character states (pleisiomorphic/apomorphic) taken from Kurzweil and Linder (1998).

Bolus (1889) dismissed the use of gynostemium characters as being indicative of natural groups, as they do not correlate with differences in perianth and vegetative characters. However, Kurzweil and Linder (1998) found that variation in rostellum structure is largely congruent with their cladogram and therefore indicative of the major phylogenetic lineages in the genus. Molecular phylogenetic data do not, however, support this interpretation: the rostellum and viscidium characters that characterize the members of *Bifidum* have arisen at least three times (Fig. 5). The spurs of *Satyrium* have played an important role in previous classifications (Lindley 1830-1840; Bolus 1889; Schlechter 1902). Thereby a distinction was made between spur type (saccate or slender) and spur length (shorter or longer than the ovary). Groups that were based on spur length (Lindley 1830-1840) are shown to be polyphyletic (Fig. 3; Table 1). Bolus (1889) considered spur length too labile to use for infrageneric classification, as he was aware of intraspecific variation. Instead, Bolus (1889) and Schlechter (1902) used spur type as a diagnostic character. Our phylogeny indicates that spur type is a conservative character with only few changes and hence useful for diagnosing clades (Fig. 5). *Satyrium muticum* Lindl. is the only "misfit" species: Lindley (1830-1840) placed it among other short-spurred species but several features, such as its robust flowers, one or two leaves closely appressed to the ground, flower color, and its rostellum lacking parallel or bifid arms (Hall 1982) reinforce its placement in the satyrium clade, indicating secondary loss of spurs. Other morphological characters that are congruent with the molecular data are possession of lateral an undulate anticlinal wall of the seed coat supporting the branch subtending the brachysaccium clade and a strongly contracted galea entrance supporting the chlorocorys clade (Fig. 5).

The present study rejects previous classifications but only provides a starting point for a new classification. Further sampling of morphology, DNA sequence data from multiple genomes, and addition of more taxa is needed to provide a more robust phylogenetic hypothesis and to find diagnostic characters to identify the groups.

ACKNOWLEDGMENTS

We thank Bill Liltved for providing plant material, as well as the Jodrell Laboratory of the Royal Botanic Gardens, Kew for providing DNA. Hubert Kurzweil provided plant material and reviewed the manuscript. Wim Calame, two anonymous reviewers, Patrick Herendeen, and Alan Meerow provided useful comments on an earlier version of the manuscript. The National Geographic society supported fieldwork. This work was supported by SNSF grant 31-66594.01, "Adaptive and rapid radiation in the Cape flora."

LITERATURE CITED

Bellstedt D. U., H. P. Linder, and E. H. Harley. 2001. Phylogenetic relationships in *Disa* based on non-coding *trnL-trnF* chloroplast sequences: evidence of numerous repeat regions. Am. J. Bot. 88: 2088-2100.

Bose T. K. and S. K. Bhattacharjee. 1980. *Orchids of India*. Calcutta: Naya Prokash,.

Chen S., Z. Tsi, and Y. Luo. 1999. *Native orchids of China in colour*. Science Press, Beijing.

Bolus H. 1889. Contribution to South African Bot.. Bot. J. Linn. Soc. 25: 156-210.

Bull J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. Syst. Biol. 42: 384-397.

Cameron K. M., M. W. Chase, W. M. Whitten, P. J. Kores, D. C. Jarrell, V. A. Albert, T. Yukawa, H. G. Gillis, and D. H. Goldman. 1999. A phylogenetic analysis of the Orchidaceae: evidence from *rbcL* nucleotide sequences. Am. J. Bot. 86: 208-224.

Cribb P. J. and S. Thomas. 1997. Orchidaceae. *Flora of Ethiopia and Eritrea Vol. 6*, eds S. Edwards, S. Demissew, and I. Hedberg. Addis Ababa and Uppsala: The National Herbarium and The Department of Systematic Botany.

De Queiroz A. 1993. For consensus (sometimes). Syst. Biol. 42: 368-372.

De Queiroz A, M. J. Donoghue, and J. Kim. 1995. Separate versus combined analysis of phylogenetic evidence. Annu. Rev. Ecol. Syst. 26: 657-681.

Douzery E. J. P., A. M. Pridgeon, P. J. Kores, H. P. Linder, H. Kurzweil, and M. W. Chase. 1999. Molecular phylogenetics of Diseae (Orchidaceae): a contribution from nuclear ribosomal ITS sequences. Am. J. Bot. 86: 887-899.

Doyle J. and J. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19: 11-15.

Dressler R. L. 1993. *Phylogeny and classification of the orchid family*. Cambridge: Cambridge University Press.

Eldenas P. and H. P. Linder. 2000. Congruence and complementarity of morphological and *trnL-trnF* sequence data and the phylogeny of the African Restionaceae. Syst. Bot. 25: 692-707.

Ellis A. G. and S. D. Johnson 1999. Do pollinators determine hybridization patterns in sympatric *Satyrium* (Orchidaceae) species? Plant Syst. Evol. 219: 137-150.

Farris J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence. Cladistics 10: 315-319.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39: 783-791.

Geerinck D. 1984. Orchidaceae (première partie). *Flore d'Afrique Centrale*. Meise : Jardin botanique national de Belgique.

Goldblatt P. 1978. An analysis of the flora of southern Africa: its characteristics, relationships, and origins. Ann. Mo. Bot. Gard. 65: 369-436.

Goldblatt P., V. Savolainen, O. Porteous, I. Sostaric, M. Powell, G. Reeves, J. C. Manning, T. G. Barraclough, and M. W. Chase. 2002. Radiation in the Cape flora and the phylogeny of peacock irises *Moraea* (Iridaceae) based on four plastid DNA regions. Mol. Phyl. Evol. 25: 341-360.

Gravendeel B., M. W. Chase, E. F. de Vogel, M. C. Roos, T. H. M. Mes, and K. Bachman. 2001. Molecular phylogeny of *Coelogyne* (Epidendroideae; Orchidaceae) based on plastid RFLPs, *matK*, and nuclear ribosomal ITS sequences: evidence for polyphyly. Am. J. Bot. 88: 1915-1927.

Hall A.V. 1982. A revision of the Southern African species of *Satyrium. Contributions from the Bolus Herbarium* 10: 1-142.

Hillis D. M. 1995. Approaches for assessing phylogenetic accuracy. Syst. Biol. 44: 3-16.

Hillis D. M., B. K. Mable, and C. Moritz. 1996. Applications of Molecular Systematics Pp. 515-543 in *Molecular Systematics*, eds. D. M. Hillis, C. Moritz and B. K. Mable. Sunderland: Sinauer Associates.

Holmgren, P. K. and W. Keuken. 1974. *Index Herbariorum, ed. 6.* Utrecht: Oosthoek, Scheltema & Holkema.

Huelsenbeck J. P., D. M. Hillis, and R. Nielsen. 1996. A likelihood-ratio test of monophyly. Syst. Biol. 45: 546-558.

Humbert H. 1939. Orchidées (tome 1). *Flore de Madagascar,* eds. H. Perrier de la Bathie. Tananarive: Imprimerie Officielle.

Kluge A. G. 1989. A concern for evidence and the phylogenetic hypothesis of relationship among *Epicrates* (Boidae, Serpentes). Syst. Zool. 38**:** 7-25.

Kocyan A., E. Conti, Y.-L. Qiu, and P. K. Endress. 2004. A phylogenetic analysis of Apostasioideae (Orchidaceae) based on ITS, *trnL-F* and *matK*. Plant Syst. Evol. 247: 203-213.

Kränzlin F. 1901. *Orchidearum genera et species*. Mayer and Müller, Berlin.

Kurzweil H. 1996. Floral morphology and ontogeny in subtribe Satyriinae (Fam. Orchidaceae). *Flora* 191: 9-28.

Kurzweil H. and H. P. Linder. 1998. A phylogenetic analysis of the genus *Satyrium* (Orchidaceae). *Beiträge zur Biologie der Pflanzen* 71: 101-181.

Kurzweil H. and H. P. Linder. 2001. Subtribe Satyriinae. Pp. 48-56 in: *Genera Orchidacearum. Volume 2 Orchidoideae (Part one)* eds. Pridgeon A. M., P. J. Cribb, M. W. Chase and F. N. Rasmussen. Oxford: Oxford University Press.

Kurzweil H., H. P. Linder, W. L. Stern, and A. M. Pridgeon. 1995. Comparative vegetative anatomy and classification of Diseae (Orchidaceae). Bot. J. Linn. Soc. 117: 171-220.

La Croix I. and P. J. Cribb. 1995. Orchidaceae. *Flora Zambesica Vol. 11 part 1*, ed. G. V. Pope. London: Flora Zambesica Managing Committee.

Lavin M. R. T. Pennington, B. B. Klitgaard, J. I. Sprent, H. Cavalcante de Lima, and P. E. Gasson. 2001. The dalbergioid legumes (Fabaceae): delimitations of a pantropical monophyletic clade. Am. J. Bot. 88: 503-533.

Linder H. P. 1986. Notes on the phylogeny of the Orchidoideae, with particular reference to the Diseae. *Lindleyana* 1: 51-64.

Linder H. P. and H. Kurzweil. 1994. The phylogeny and classification of the Diseae (Orchidoideae: Orchidaceae). Ann. Mo. Bot. Gard. 81: 687-713.

Linder H. P. and H. Kurzweil. 1999. *Orchids of Southern Africa*. Rotterdam: Balkema.

Lindley J. 1830-1840. *The genera and species of Orchidaceaous plants*. London: Ridgways.

Miyamoto M. M. and W. M. Fitch. 1995. Testing species phylogenies and phylogenetic methods with congruence. Syst. Biol. 44: 64-76.

Moody M. L., L. Hufford, D. E. Soltis, and P. E. Soltis. 2001. Phylogenetic relationships of Loasaceae subfamily Gronovioideae inferred from *matK* and ITS sequence data. Am. J. Bot. 88: 326-336.

Nixon K. C. and J. M. Carpenter. 1996. On simultaneous analysis. Cladistics 12: 221-241.

Polunin O. and A. Stainton. 1984. *Flowers of the Himalaya*. Delhi: Sri Aurobindo Ashram Press.

Posada D. and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Pridgeon A. M., R. Solano, and M. W. Chase. 2001. Phylogenetic relationships in Pleurothallidinae (Orchidaceae): combined evidence from nuclear and plastid DNA sequences. Am. J. Bot. 88: 2286-2308.

Rambaut A. and N. C. Grassly. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Computational and Applied Biosciences 13: 235-238.

Reeves G., M. W. Chase, P. Goldblatt, P. Rudall, M. F. Fay, A. V. Cox, B. Lejeune, and T. Souza-Chies. 2001. Molecular Syst.s of Iridaceae: evidence from four plastid DNA regions. Am. J. Bot. 88: 2074-2087.

Ronquist F. and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Salazar G. A., M. W. Chase, M. A. Soto Arenas, and M. Ingrouille. 2003. Phylogenetics of Cranichideae with emphasis on Spiranthinae (Orchidaceae, Orchidoideae): evidence from plastid and nuclear DNA sequences. Am. J. Bot. 90: 777-795.

Schlechter R. 1902. Monographie der Diseae. Bot. Jahrb. Syst. Pflanzengesch. Pflanzengeogr. 31: 134-313.

Schlechter R. 1926. Das System der Orchidaceen. Notizblatt des Botanischen Gartens und Museums zu Berlin-Dahlem 9: 563-591.

Simmons M. P. and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49: 369-381.

Summerhayes V. S. 1968a. Orchidaceae (part 1). *Flora of tropical East Africa,* eds. E. Milne-Redhead and R. M. Polhill. London: Crown Agents for Oversea Governments and Adminstrations.

Summerhayes V. S. 1968b. Orchidaceae Vol III Part I. *Flora of west tropical Africa,* eds. J. Hutchinson and J. M. Dalziel (first ed.) and F. N. Hepper (second ed.). London: Crown Agents for Oversea Governments and Adminstrations.

Swofford D. L. 2000. PAUP*: Phylogenetic analysis using parsimony, version 4.0b. Sunderland: Sinauer Associates.

Taberlet P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105-1109.

Thompson J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Van der Heede C. J., R. L. L. Viane, and M. W. Chase. 2003. Phylogenetic analysis of *Asplenium* subgenus *Ceterach* (Pteridophyta: Aspleniaceae) based on plastid and nuclear ribosomal ITS DNA sequences. Am. J. Bot. 90: 481-495.

Wendel J. F. and J. J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. Pp. 265-296 in *Molecular Syst.s of plants II: DNA sequencing*, eds. D. E. Soltis, P. S. Soltis, and J. J. Doyle. Norwell: Kluwer Academic Publishers.

White T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315-322 in *PCR protocols: a guide to methods and applications*, eds. M. Innis, D. Gelfand, J. Sninsky, and T. White. San Diego: Academic Press.

Whitten W. M., N. H. Williams, and M. W. Chase. 2000. Subtribal and generic relationships of Maillarieae (Orchidaceae) with emphasis on Stanhopeinae: combined molecular evidence. Am. J. Bot. 81: 1842-1856.

Wood J. R. I. 1997. *A handbood of the Yemen flora*. Kew: Royal Botanic Gardens.

Yoder A. D., J. A. Irwin, and B. A. Payseur. 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. Syst. Biol. 50: 408-424.

Young N. D. and J. Healy. 2003. Gapcoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4: 6

Dealing with Incongruence in the Quest for the Species Tree: A Case Study from the Orchid Genus *Satyrium*

TIMOTHEÜS VAN DER NIET[1] AND H. PETER LINDER[1]

[1]*Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008, Zurich, Switzerland; E-mail: niet@systbot.unizh.ch (T.N.)*

Submitted to Systematic Biology

Abstract.—A species tree was reconstructed for *Satyrium*, a mainly African terrestrial orchid genus, which is known to hybridize in nature. Phylogenetic analysis of both plastid and ribosomal nuclear DNA sequences for 63 species, revealed extensive topological conflict when these two genomic data partitions were analysed separately. Here we describe a detailed protocol to deal with incongruence involving three steps: identifying incongruence, assessing the cause of incongruence, and reconstructing the species tree. The Incongruence Length Difference test revealed that many cases of incongruence were in fact, non-significant. For the remaining significant cases, results from taxon jack-knife experiments and parametric bootstrap suggested that non-biological artefacts such as sparse taxon sampling and long-branch attraction could be excluded as causes for the observed incongruence. In order to evaluate biological causes, such as orthology/paralogy conflation, lineage sorting, and hybridization, the number of events was counted that needs to be invoked *a-posteriori* to explain the observed pattern. In most cases where incongruence was significant, this resulted in an almost equal number of events for each of these different causes. Only for the three species from south east Asia, that form a monophyletic clade, hybridization was favoured over the alternative causes. This conclusion is based on the large number of events that needs to be invoked, in order for either orthology/paralogy conflation or lineage sorting to have been the cause of the incongruence. Morphological evidence further supports a hybrid origin of this clade. The final species tree presented here is the product of the combined analysis of both plastid and ribosomal nuclear DNA sequences for all non-incongruent species and *a-posteriori* grafting of the incongruent clades or accessions onto the tree. This tree provides the best phylogenetic hypothesis to date, and serves as a template for subsequent evolutionary studies.

Keywords: incongruence; ILD test; ITS; hybridization; plastid; *Satyrium*; species tree

INTRODUCTION

Phylogenetic species trees are a necessary ingredient of any study of speciation (e.g. Barraclough and Nee 2001). To reconstruct a species tree, firstly DNA sequences of single genes from exemplar specimens are used to reconstruct gene trees (Soltis et al. 1998). Once sufficient genes from different genomes have been sampled, a species tree can be inferred (Doyle 1992, 1997; Maddison 1997). Multiple gene trees, however, may be incongruent. In fact, the more genes are sampled, the more incongruence among gene trees is found (e.g. Rokas et al. 2003).

Two classes of incongruence can be distinguished. Firstly, non-biological artefacts such as insufficient taxon sampling (e.g. Stockley et al. 2005) and long-branch attraction (e.g. Felsenstein 1978; Sanderson et al. 2000; Fan and Xiang 2003; Kennedy et al. 2005) can cause incongruence. This may happen even if all data partitions share the same underlying phylogenetic history. Secondly, incongruent gene trees may in fact depict phylogenetic relationships correctly as a result of different underlying phylogenetic histories (Doyle 1992; Maddison 1997). Several biological processes such as orthology/paralogy conflation (e.g. Vanderpoorten et al. 2004), horizontal gene transfer (e.g. Won and Renner 2003), lineage sorting (e.g. Doyle et al. 2004), and hybridization (e.g. Rieseberg et al. 1996) may lead to different phylogenetic histories among genes, and hence incongruent gene trees (Wendel and Doyle 1998).

In order to find the species tree for a particular group, we need to determine whether or not the incongruent nodes of gene trees should be considered part of a highly reticulate species tree. For this we need to establish whether any incongruent node is the result of non-biological artefacts or biological phenomena below the species level that are irrelevant to the species tree, or whether it is the result of hybridization.

Even though it is almost impossible to exclude any potential cause of incongruence *a-priori* with only a set of gene trees at hand, only rarely are multiple causes considered and carefully examined. More often, incongruent clades or accessions are either included in a combined analysis without paying further attention to the incongruence (e.g. Wang et al. 2004) or they are left out of the analysis completely (e.g. Alejandro et al. 2005; Kyndt et al. 2005). Alternatively, only a subset of possible causes of incongruence is considered (e.g. Lantz and Bremer 2005). Examples of studies that have carefully and explicitly tested multiple causes of incongruence are rare (e.g. Wiens and Hollingsworth 2000; Duvall and Ervin 2004). Both of these studies were carried out above the species-level and therefore have limited relevance to the reconstruction of a species tree. Here we demonstrate that detailed

assessment of potential causes of incongruence can contribute to better resolution of the species tree.

We reconstruct a species tree of the terrestrial orchid genus *Satyrium* using DNA sequence data from the plastid and nuclear genome. The 90 species of *Satyrium* are distributed throughout sub-Saharan Africa and Madagascar, with four species extending into Asia (Summerhayes 1968a; Summerhayes 1968b; Bose and Bhattacharjee 1980; Geerinck 1984; Polunin and Stainton 1984; la Croix and Cribb 1995; Cribb and Thomas 1997; Wood 1997; Chen et al. 1999; Linder and Kurzweil 1999, Van der Niet and Cribb submitted). Several features such as distribution range, diverse morphology (Hall 1982; Kurzweil and Linder 1998) and presence of several pollination systems (Johnson 1997; Johnson and Steiner 1999) make *Satyrium* an attractive group to address questions related to speciation. A previously published phylogeny based on limited sampling of DNA sequences from the nuclear and plastid genome of 39 species revealed several cases of incongruence (Van der Niet et al. 2005). Hybridization is reported to occur between several species of *Satyrium*, and even between species assigned to different sections of the genus (Hall 1982; Ellis and Johnson 1999).

Given the previously observed incongruence and the presence of hybridization in the genus, we address the following questions using a dataset with both increased taxon and character sampling: (1) Can we identify any well-supported cases of incongruence? (2) Is the observed incongruence the result of non-biological artifacts such as insufficient taxon sampling or long-branch attraction? or (3) can we assign the incongruence to biological phenomena such as orthology/paralogy conflation, lineage sorting and hybridization? Finally (4) what is the best species tree for *Satyrium*?

MATERIAL AND METHODS

Taxon Sampling

We included 63 of the 90 species of *Satyrium* (ca. 70% sampling), representing both the geographical and morphological range of the genus. In cases where it was deemed potentially informative given the geographical distribution, flowering time, and morphological diversity of a species, multiple accessions were sampled. In four cases both subspecies and varieties of a species were sampled. Nine outgroups were included representing the major lineages of Orchideae and Diseae. All voucher information and genebank accession numbers can be found in Appendix 1.

Molecular Methods

Molecular methods followed Van der Niet et al. (2005) unless otherwise indicated. The *trnL* intron and *trnL-F* intergenic spacer (hereafter, trnLF) were amplified using the same primers as Van der Niet et al. (2005) The small part of the plastid *matK* gene and *trnK* intron previously sequenced for Van der Niet et al. (2005) was extended to include the entire coding region of *matK* and almost the entire 5' end of the *trnK* intron (hereafter, matK). The external primers -19F (Kores et al. 2000) and R1 (Kocyan et al. 2004) were used for the PCR. For sequencing, and in case the PCR of the entire region failed using the external primers, internal primers 580F, 596R, 1082F, and 1361R (Kocyan et al. 2004) were used. The plastid *trnS-G* intergenic spacer (hereafter, trnSG) was amplified using primers trnS and trnG (Hamilton 1999). Both matK and trnSG were amplified on a Biometra® T1 thermocycler using the program: 3 minutes denaturation at 95°C followed by 35 cycles of 30 seconds denaturation at 95°C, 1 minute annealing at 52°C, 1 minute extension at 72°C. The program was terminated with 4 minutes final extension at 72°C. Amplification of ITS1, the 5.8S gene, and ITS2 (hereafter, ITS) was performed using the same primers and amplification protocol as Van der Niet et al. (2005). In cases where direct sequencing resulted in ambiguous electropherograms, the PCR product was cloned using chemically competent *Escherichia coli* of a TOPO® TA Cloning kit (Invitrogen, Carlsbad, California) following the manufacturers protocol. Four colonies were selected and amplified using the M13 forward and M13 reverse primers (Invitrogen, Carlsbad, California) on a Biometra® T1 thermocycler using the program: 5 minutes denaturation at 95°C followed by 34 cycles of 1 minute denaturation at 94°C, 1 minute annealing at 55°C, and 90 seconds extension at 72°C. The program was terminated with 7 minutes final extension at 72°C.

Data Matrix Composition

Editing and alignment followed Van der Niet et al. (2005). The sequences of trnLF and trnSG were trimmed by excluding the *trn* gene sequences, such that these sequences consisted of intron and spacer sequences only. A one basepair (bp) insertion in matK for *Habenaria* that would have disrupted the reading frame for all other species was excluded from further analyses, as well as a 10 bp inversion in trnSG. Polymorphism of ITS that did not require cloning for readability was coded using the IUPAC coding. Unambiguously aligned gaps were coded under the criteria of 'simple indel coding' (Simmons and Ochoterena 2000) using Gapcoder (Young and Healy 2003). Three data matrices were composed. (1) The plastid matrix included the sequences of all plastid loci and coded gaps for 91 taxa (all

Satyrium accessions + outgroups). (2) The ITS matrix included the nuclear ITS sequences of all *Satyrium* accessions. Clones that formed a monophyletic clade in an initial analysis were only represented by one clonal sequence. In case that the clones of an accession were not monophyletic in an initial analysis, clones with different phylogenetic affinities were all included in the ITS matrix. Gaps were not coded for ITS due to difficult alignment. This matrix does not include any outgroups due to the fact that it was impossible to unambiguously align these with the *Satyrium* sequences. (3) The combined matrix included both the plastid and nuclear loci. In an attempt for completely overlapping taxon sampling in this matrix, all outgroups were removed and the root position was enforced according to the root found for the plastid analyses. Accessions with several ITS clones that were not monophyletic were represented by a clonal sequence that was congruent with the plastid topology.

Phylogenetic Analyses

The plastid and ITS matrices were analysed separately, using both parsimony and Bayesian inference. Each individual plastid locus (trnLF, trnSG, and matK) was also analysed separately. For parsimony, all characters were equally weighted and treated as unordered. Heuristic searches using PAUP* 4.0b (Swofford 2000) were started with 1000 stepwise random addition sequence replicates (RASR), holding ten trees at each step, followed by TBR swapping, saving maximally 100 most parsimonious trees (MPTs). All shortest trees retained in memory were used as starting trees for a second round of searching involving exhaustive TBR swapping. Character support for each node was inferred using the bootstrap procedure (Felsenstein 1985). Five hundred bootstrap replicates, each obtained from 250 stepwise RASR with three trees held each step and saving no more than ten trees, were performed**.**

We tested for saturation of substitutions by plotting the observed number of changes (the differences between two aligned sequences) against the inferred number of changes (the optimized branch lengths between two terminals) using the patristic distance table from PAUP* 4.0b (Swofford 2000). We performed the optimization for several exemplar taxa onto one of the MPTs resulting from a heuristic search of the combined matrix using ACCTRAN optimization.

For Bayesian analyses, the plastid and ITS matrices were divided according to several different partitioning schemes. These represented biological categories (e.g. intron, spacer, gene, etc.) or contiguous blocks of nucleotides (trnLF, trnSG, matK, etc.). For each of these, the optimal model of sequence evolution was calculated on a randomly chosen MPT from the set of obtained trees from parsimony analysis, using the Akaike Information Criterion (AIC)

criterion as implemented in Modeltest 3.06 (Posada and Crandall 1998). Gaps were analyzed using a maximum likelihood Markov model for variable characters only (Lewis 2001). Bayesian analyses were carried out using MrBayes v3.0B4 (Ronquist and Huelsenbeck 2003). One million generations were run with a sample taken every 1000 generations and with the first 250 samples considered as burnin. The posterior probability (PP) of each clade was calculated using this sample and expressed as percentage. To assess what partitioning scheme is optimal for the nuclear and plastid datasets, the harmonic mean values of the likelihood, which are part of the output of MrBayes, were compared (Nylander et al. 2004).

Disperis villosa was used for rooting trees resulting from analyses of the plastid matrix. The root for all trees resulting from analyses of the ITS and combined matrices was forced so as to be in the same position as the root of the plastid trees. This root position was exactly the same as was found for previously published trees that used nuclear ITS sequences (e.g. Douzery et al. 1999). The data matrices and the phylogenetic trees are deposited in Treebase study number SN2752 (http://www.treebase.org).

Identification of Incongruence

As there was no significant incongruence among plastid loci, the remainder of the paper will focus on incongruence between topologies reconstructed from the plastid and ITS datasets. To identify which accessions or clades are incongruent between the plastid and ITS topologies, strict consensus topologies from separate parsimony analyses were compared. An accession or clade was identified as incongruent if removal of that accession or clade solved the incongruence. This systematic removal of as few accessions or clades as possible is considered to be the optimal solution for solving such incongruences. If removal of two different clades equally solved the incongruence, the clade containing the least number of accessions was removed. In cases of complex incongruence, different sets of the same number of accessions or clades could be removed to solve the incongruence. All these solutions were subsequently taken into account when testing the significance of incongruence. In cases where the incongruence involved only three accessions or clades, removal of any of these will solve the incongruence. In all cases where this applied, significance of incongruence was only assessed for one of the three accessions or clades.

Incongruence Testing

Congruence among datasets was tested with the Incongruence Length Difference (ILD) test (Farris et al. 1994), implemented as Partition Homogeneity Test in PAUP* 4.0b

(Swofford 2000). Given that the primary focus of this study is a species tree for *Satyrium*, outgroups were always excluded while running the ILD. We used a similar approach as Van der Niet et al. (2005) where all accessions and clades causing the ILD test to return a significant value in the combined matrix were pruned. Subsequently, these incongruent accessions or clades were individually added to the pruned combined matrix and the ILD test was repeated. One hundred randomizations were run with 50 RASR, holding two trees each step, and saving no more than five trees. If the P-value was close to the significance level (determined here as P=0.05), an additional round was run with 1000 randomizations and the same search strategy. Only significantly incongruent cases were considered for the subsequent tests for this paper.

The support for conflicting positions of incongruent accessions or clades was assessed with 100 bootstrap replicates, with each replicate calculated from 50 RASR, holding three trees each step and and saving no more than five trees. These bootstrap analyses were carried out for both the plastid and ITS matrices respectively, each time containing only one incongruent accession or clade similar to the taxon sampling procedure used for the ILD test. The bootstrap value supporting the placement of an incongruent accession or clade of the least supported dataset was plotted against the P-value of the ILD test. If more than one node defined the incongruence, the highest bootstrap value among these multiple nodes was used. The critical bootstrap value at which the ILD test returns a significant result was thereby determined. This critical value was used in subsequent analyses as a quick proxy to determine statistically significant incongruence.

Taxon Sampling

If incongruence is the result of sparse taxon sampling, further reduction of the taxon sampling should result in an increase of cases of significant incongruence. To test for this we applied taxon jack-knifing by randomly removing 20%, 40%, 60% and 80% of the taxa. For each jack-knife level we generated ten plastid and ITS data matrices, which were analysed separately using the bootstrap. For each matrix 100 bootstrap replicates were run with 50 RASR per bootstrap replicate, holding three trees each step, and saving no more than five trees. We counted the number of cases of incongruence between the plastid and ITS matrices, as identified by the critical bootstrap value described above. We plotted this number averaged over the ten experiments against the level of taxon sampling.

Long-branch Attraction

We tested whether the incongruence was caused by long-branch attraction in one of the two datasets by calculating the difference in steps between the MPT for a dataset, and the length obtained for that dataset if the topology was constrained to the MPT of the alternative dataset. To test the statistical significance of this value, we compared it to a null distribution generated by simulation through parametric bootstrap (Huelsenbeck et al. 1996). The null distribution is generated under the hypothesis that the difference in the number of steps between a constrained topology and the MPT is caused by the stochastic process of sequence evolution. If this resulted in the presence of long branches in the simulation, then the most parsimonious unconstrained topology (including the incorrectly inferred clades that result from long-branch attraction) will have a shorter tree length than the constrained topology onto which the sequences were simulated. The simulations were done using pruned matrices that only contained one significantly incongruent accession or clade at a time. This method is valid because the original assessment of incongruence was done on similar matrices. In cases of complex incongruence we only carried out the analyses for one subset of incongruent accessions or clades. The plastid and ITS matrices were analysed using parsimony in PAUP* 4.0b (Swofford 2000). For simulating ITS data, we selected one randomly chosen MPT from phylogenetic analysis of the plastid dataset. This topology was used to calculate the parameters of the model of ITS sequence evolution using the AIC implemented in Modeltest 3.06 (Posada and Crandall 1998). Subsequently, given these values and the plastid topology, branch lengths were calculated from the ITS matrix using Maximum Likelihood in PAUP* 4.0b (Swofford 2000). The topology, branch lengths and sequence parameters were used as input for Seq-Gen.v1.2.6 (Rambaut and Grassly 1997) to generate 100 simulated ITS datasets. The number of characters that were simulated for a sequence was calculated as follows: (aligned sequence length) – (total number of characters caused by gaps and missing data/number of taxa). The 100 simulated datasets were analysed with and without the original topological constraint of the incongruent accession or clade involved, using parsimony. Each search included ten RASR, holding three trees each step and saving no more than five trees. Although this search strategy does not necessarily find all MPTs, it will most likely find the length of the MPT, which is the variable of interest. The difference in steps between constrained and unconstrained analysis was plotted in a frequency diagram. For simulating plastid data, the same procedure was reciprocally carried out. The observed value was tested against this null distribution for significance ($\alpha<0.01$).

Orthology/paralogy Conflation, Lineage Sorting, and Hybridization

For orthology/paralogy conflation, lineage sorting, and hybridization, the minimum number of events that needs to be postulated to arrive at the observed pattern was counted. Specifically, for orthology/paralogy conflation the minimum number of gene duplications and extinctions of copies was counted, asssuming that the PCR would result in amplification of all present copies. This procedure was only done for ITS because there is no reason to believe gene duplication has occurred for plastid loci (Soltis et al. 1998). For lineage sorting the minimum number of polymorphisms arising and cladogenic events that these polymorphisms must have persisted, was counted. The calculations for both orthology/paralogy conflation and lineage sorting were done for each significantly incongruent accession or clade individually, while ignoring the other accessions or clades that were significantly incongruent, their positions being essentially ambiguous. For hybridization, we established whether dispersal and extinction events need to be postulated to make hybridization between putative parental accessions or clades possible. Specifically, we assumed that the hypothetical putative ancestral parental accessions or clades had either a sympatric or parapatric distribution (as inferred from present distribution ranges).

For ITS clones of the same accession that were not monophyletic, the minimum number of gene duplications and extinctions of copies that need to be postulated were counted. The minimum number of extinction events that needs to be postulated was also calculated if hybridization was the cause of the presence of multiple copies.

Reconstruction of the Species Tree for Satyrium

A species tree was reconstructed as follows: significantly incongruent accessions or clades were pruned from the combined matrix. The pruned combined matrix was analysed using both parsimony and Bayesian inference, and support was assessed using the bootstrap and Bayesian PP. All search strategies were as described above for the separate analyses of the plastid and ITS dataset. The significantly incongruent accessions or clades were *a-posteriori* inserted onto the tree topology that resulted from the combined analysis. The accessions and clades were attached at the two different positions in the topology that reflect the plastid and nuclear phylogenetic relationships retrieved from the separate anlyses of the plastid and nuclear datasets. For incongruent clades consisting of more than two accessions, a combined phylogenetic analysis, using parsimony with a branch and bound algorithm, was carried out for that clade. The whole clade, with the relationships so obtained, was inserted

into its appropriate positions. For the tree presented here, all multiple accessions of one species that were monophyletic were represented as one terminal.

RESULTS

Amplification, Sequencing, and Alignment

Amplification, sequencing and alignment were unproblematic for all plastid loci. PCR amplification for ITS was equally straightforward. For a few accessions the electropherograms resulting from direct sequencing of PCR products of ITS showed evidence of extensive polymorphism. For those accessions, the PCR product was cloned. Not all of the sequences of an accession generated through cloning were monophyletic in an initial phylogenetic analysis. Characteristics of the three datasets can be found in Table 1. The ITS matrix contains more parsimony informative characters than the plastid matrix, even though the number of aligned bases is about 5-fold higher for the total plastid matrix.

Phylogenetic Analyses

Parsimony analyses of the plastid and ITS dataset, resulted in generally well-supported cladograms (Figs.1 and 2, respectively). The percentage of nodes supported by 80%-100% Bootstrap Support (BS) was lower for the plastid dataset as compared to the ITS dataset (Table 1). However, when all support ranging from 50%-100% BS was taken into account, this percentage was higher for the plastid dataset.

The number of inferred vs. observed substitutions calculated for several exemplar taxa onto one of the MPTs of a combined phylogenetic analysis was plotted (Fig. 3). There is a stronger deviation from a perfect correlation for the ITS than the plastid dataset.

For Bayesian analysis of the plastid data, the partitioning scheme assigning different models of sequence evolution to different codon positions for the coding part of matK, the introns, intergenic spacers, and gaps, returned the highest value of the harmonic mean (results not shown). This harmonic mean value was higher than that obtained for a partitioning scheme that further divided the introns into the categories of *trnL-F* intron and *trnK* intron, and spacer into *trnL-F* spacer and *trnS-G* spacer. For the ITS dataset the harmonic mean score of Bayesian analysis was highest for a partitioning scheme using three partitions corresponding to ITS1, the 5.8S gene, and ITS2. The topologies resulting from different partitioning schemes were all very similar (results not shown). The Bayesian 95% majority rule consensus trees were congruent with the strict consensus trees from parsimony analyses. Here we present only the PP values above 95% plotted onto the strict consensus trees from

TABLE 1. Sequence characteristics of all sequence datasets used in this study, as well as the combined matrices. Calculations of support values for the plastid tree were only carried out for ingroup nodes, whereas the analyses included full taxon sampling. In all other cases calculations were completed for the taxon sampling given in brackets.

partition	sequence length ingroup	aligned length (including outgroup)	number of autapomorphies ingroup	PIC[1] ingroup	% ingroup nodes supported by 80-100% BS	% ingroup nodes supported by 50-79% BS	% ingroup nodes supported by 50-100% BS	% ingroup nodes supported by ≥95% PP
matK 1st codon position		520	37	32				
matK 2nd codon position		520	21	31				
matK 3rd codon position		520	42	50				
matK all codon positions	1488-1560	1560	100	113				
matK gaps		8	4	2				
trnK intron nucleotides	262-300	349	14	23				
trnK intron gaps		16	5	4				
matK total	1770-1844	1933	123	142	34	26	60	
trnSG nucleotides	550-633	788	55	74				
trnSG gaps		50	17	11				
trnSG total					21	17	38	
trnLF intron nucleotides	261-434	702	9	20				
trnLF spacer nucleotides	310-326	354	18	24				
trnLF gaps		42	14	10				
trnLF total	579-758				16	23	39	
plastid total (91 taxa)			459	536	54	23	77	73
its excl clones (82 taxa)	651-684	720	83	366	59	17	77	69
combined (62 taxa)					67	18	85	79

[1]PIC = Parsimony Informative Characters

parsimony analyses, as these strict consensus trees were finally used for the assessment of incongruence (Figs.1 and 2). The percentage of nodes with a PP of ≥95% is higher for the plastid dataset compared to the ITS dataset.

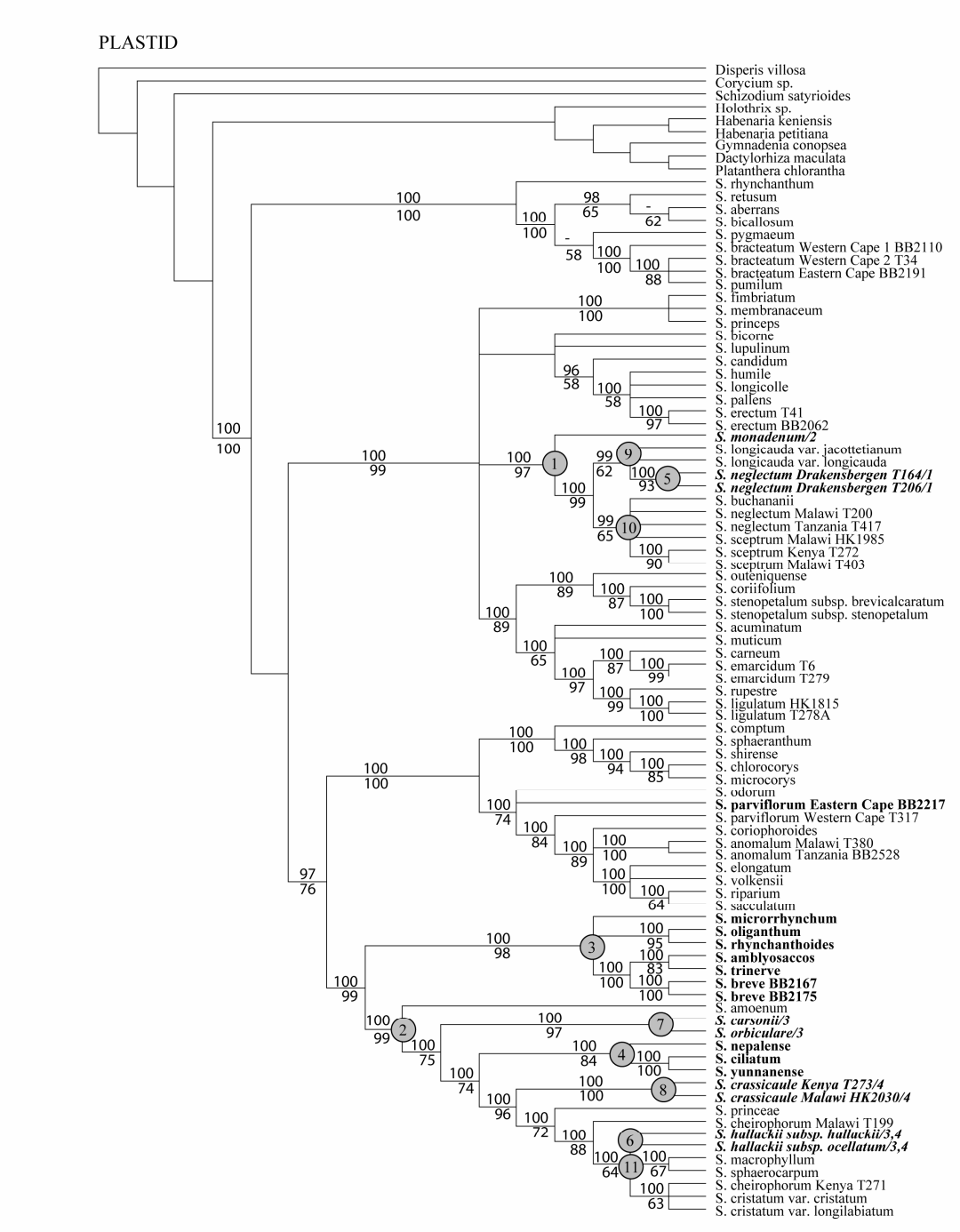


FIG. 1. Strict consensus of all most parsimonious trees resulting from individual analyses of the plastid dataset. Values above the branches are Bayesian Posterior probability (only those >95% represented), values below the branches are bootstrap percentages. Accessions or clades in bold were found to be significantly incongruent. Accessions or clades in bold and italics are members of a set of accessions or clades that need to be removed to solve the incongruence in their clade. Membership of a particular set as referred to in Table 2 is indicated after the slash. Numbered clades are referred to in the text. If multiple accessions of the same species were included, they are differentiated by their geographical origin and/or accession number.

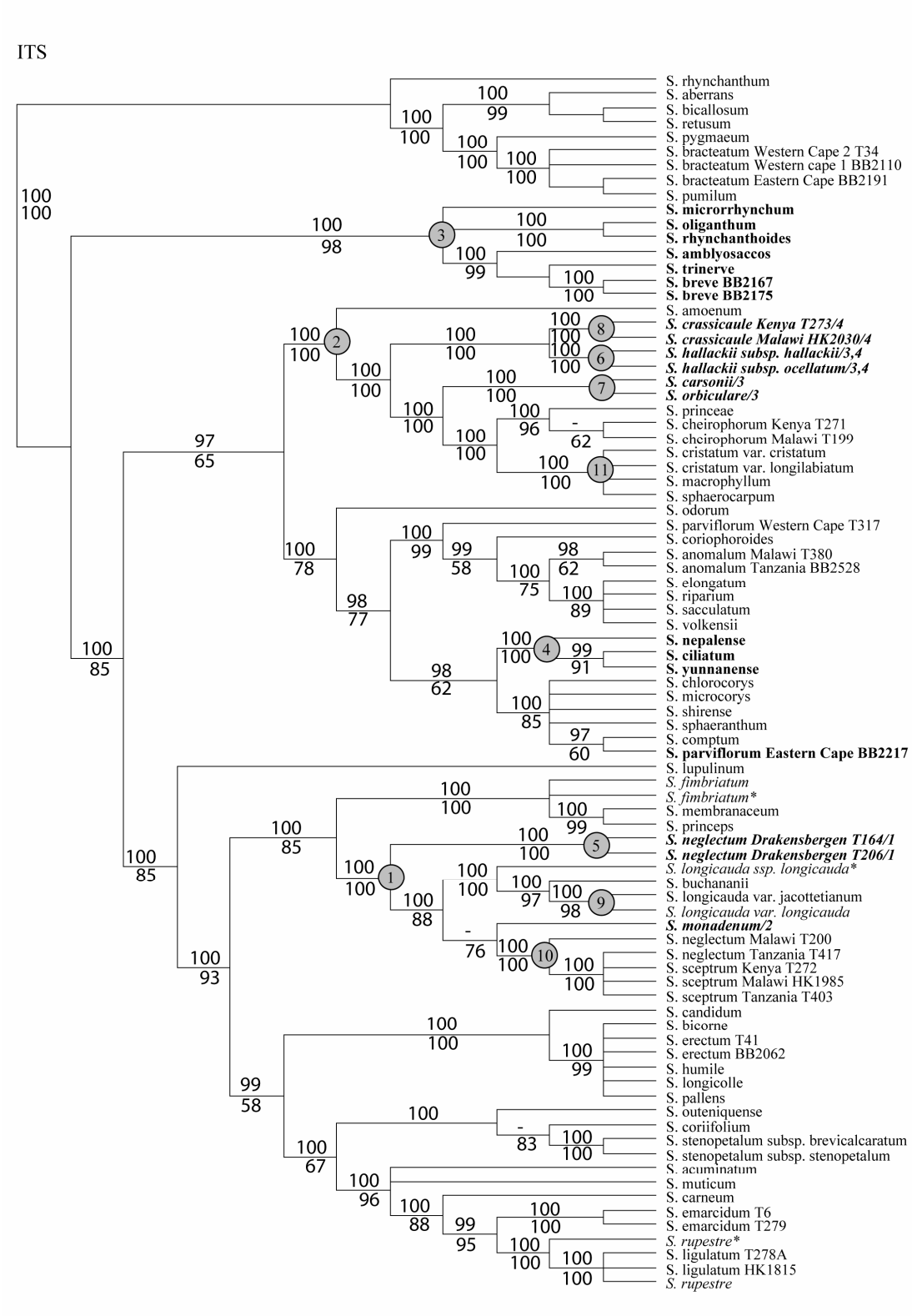


FIG. 2. Strict consensus of all most parsimonious trees resulting from individual analyses of the ITS dataset. Taxa in italics are clones, those marked with * were left out of all analyses involving the combined matrix. Accessions or clades in bold were found to be significantly incongruent. Accessions or clades in bold and italics are members of a set of accessions or clades that need to be removed to solve the incongruence in their clade. Membership of a particular set as referred to in Table 2 is indicated after the slash. Numbered clades are referred to in the text. If multiple accessions of the same species were included, they are differentiated by their geographical origin and/or accession number.

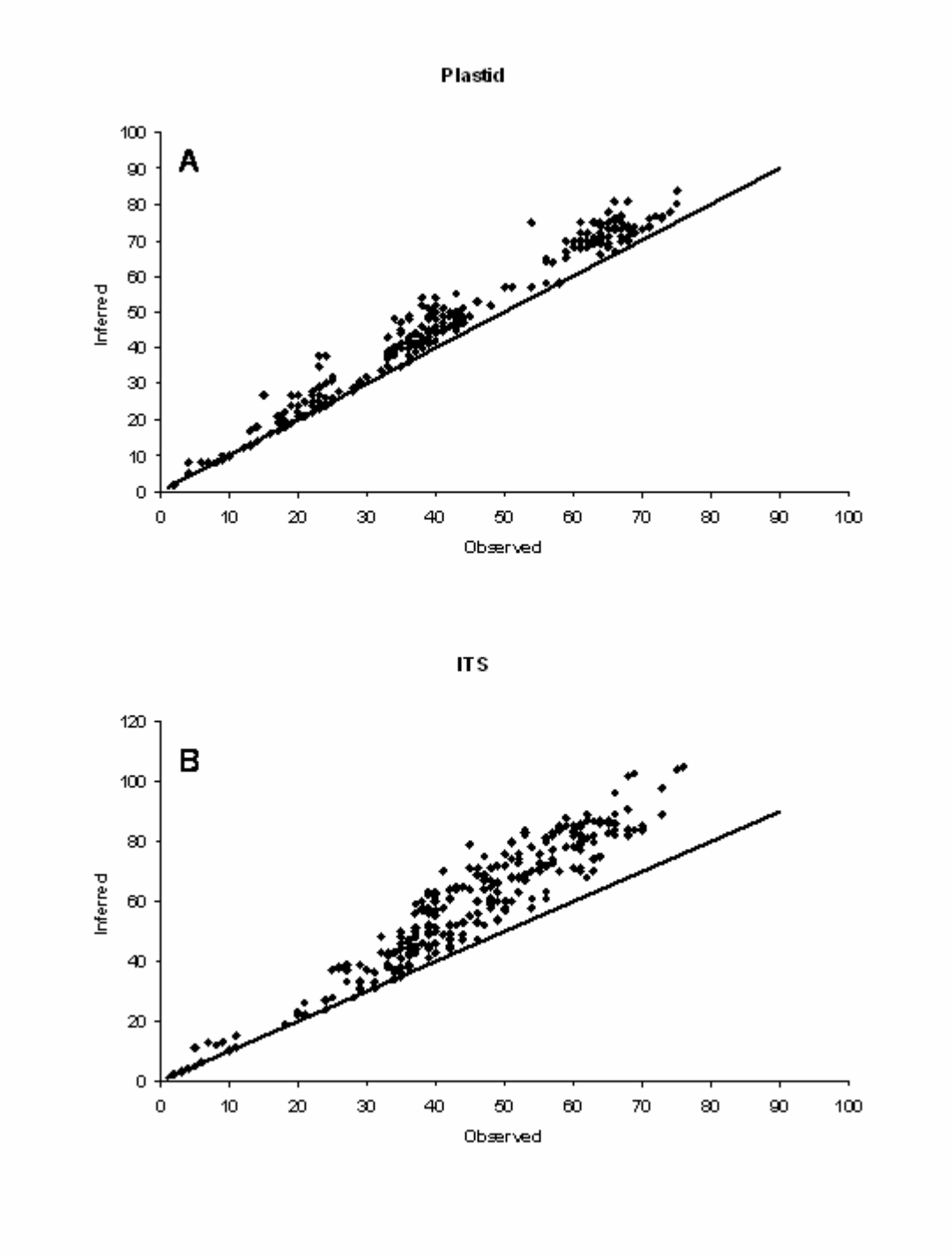


FIG. 3. Inferred changes vs. observed changes for some exemplar accessions for the plastid (a) and ITS (b) dataset. The line represents a perfect correlation with a slope of 1.

Incongruence Testing

Visual inspection of the plastid and ITS topologies (Figs.1 and 2) revealed many cases of incongruence. For most of these cases it was straightforward to determine which accession or clade should be removed to solve the incongruence. Clades 1 and 2 (as numbered in Figs.1 and 2) include complex cases of incongruence. Clade 3, that is itself incongruent, contains a nested case of incongruence. This case was evaluated within the context of clade 3 only.

TABLE 2. Accessions or clades that were incongruent based on visual inspection of the strict consensus tree from parsimony analysis of plastid and nuclear datasets, respectively. Clade numbers refer to Figs.1 and 2. Accessions and clades marked with ° are part of a case where incongruence can be reduced to involving three taxa only and where removal of any of these three taxa solves the incongruence. Accessions and clades in bold represent the nested case of incongruence that was considered only within clade 3. Accessions and clades in italics are the multiple solutions that would solve the incongruence in clade 1. Accessions and clades in bold and italics are the multiple solutions that would solve the incongruence in clade 2. If multiple solutions are possible to solve the incongruence, the respective sets of accessions and clades that are to be removed are numbered after the slash with the same number. A P-value < 0.05 is printed in bold. P-values marked with an * are the results of 1000 ILD replicates instead of the default value 100 because the value returned by 100 replicates was close to significance. The bootstrap value (>50%) that supports the incongruent positions in plastid and nuclear topology respectively results from bootstrap analysis on the pruned matrix with only the particular taxon added. If more than one value is given, it means that more than one node separates the incongruent positions. The BS in bold is the critical value that needs to be overcome in order to solve the incongruence.

Accession or clade	P-value ILD	BS plastid	BS nuclear
S. bicorne	0.24	**65**	98
S. trinerve	0.15*	85	**63**
S.parviflorum Eastern Cape	0.02*	86, 100	70, **97**
Clade 4	0.01	92, 99, **100**	51, 96, **100**
S. carneum°	0.34	**85**	99
S. odorum	0.31	87	**81**
S. aberrans°	0.3	**65**	70
Clade 3	0.015*	88, 98	62, **96**
S.lupulinum	0.56	**50**	72, 97
Clade 10/1	0.08*	58. 99	**62**
Clade 5/1	0.017*	64, **99**	88, 100
S. monadenum/2	0.016*	66, 99	70, **82**
Clade 9/2	0.187*	60, **69**	87, 100
Clade 6[1]	0.03*	74, **84**	100, 100
S. cheirophorum Kenya + Clade 11/1	0.017*	64, 69, **89**	66, 97, 100
Clade 8/1	0.08*	**97**	98
Clade 6/2	0.005*	86, **100**	96, **100**
S. cheirophorum Kenya + Clade 11/2	0.049*	59, 60, **85**	60, 94, 99
Clade 7/2	0.08*	100	**91**
Clade 6/3	0.02*	55, **94**	100, 100
S. cheirophorum Kenya/3	0.08*	58, **59**	99, 100
S. princeae/3	0.11*	**91**	99
Clade 7/3	0.047*	97	**90**
Clade 6/4	0.007*	72, **100**	96, **100**, **100**
S. cheirophorum Kenya/4	0.19*	68, **74**	99, 100
S. princeae/4	0.21*	**89**	93
Clade 8/4	0.047*	98	**92**

Table 2 highlights the incongruent accessions or clades with the P-value that was returned by the ILD test after they were individually added to a matrix that contained only accessions that were congruent between the plastid and ITS dataset. In each case where a pruned congruent matrix contained a different set of accessions, a control ILD was run which then always returned a non-significant result, as predicted. Finally, the following incongruent accessions or clades in cases of non-complex incongruence were considered for all subsequent analyses based on their significant P-value: *S. parviflorum* Eastern Cape, clade 3, and clade 4. For solving the incongruence in clade 1, both individual removal of the clade 5 and *S. monadenum* were considered. For solving incongruence in clade 2 removal of either clade 6 + clade 7 (set 3), or clade 6 + clade 8 (set 4) solved the incongruence. Both these equally parsimonious options are considered.

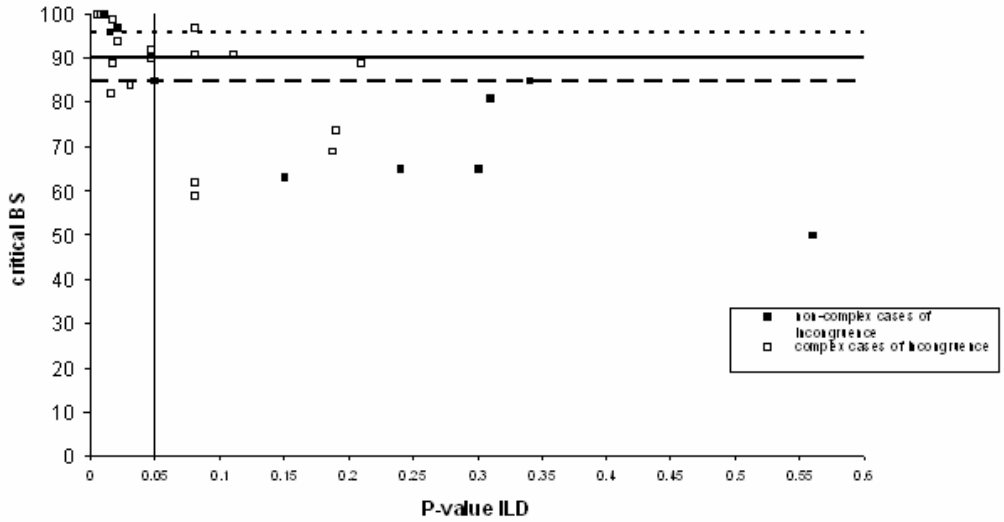


FIG. 4. Relationship between the critical bootstrap value and the P-value returned from the ILD test. Non-complex cases of incongruence are represented by filled boxes, complex cases of incongruence are represented by open boxes. The horizontal line is at 90% BS, the vertical line at P=0.05. The line with long dashes represents the lower threshold (86% BS) below which incongruence is never significant for non-complex cases of incongruence, the line with short dashes represents the upper threshold (96%) above which significance is always significant for non-complex cases of incongruence.

Plotting of the P-value returned by the ILD test against the critical bootstrap value that needs to be exceeded in order to solve the incongruence for the non-complex cases of incongruence, showed that above 96% BS incongruence was always significant. Below 85% BS, incongruence was consistently non-significant (Fig. 4). This pattern is less clear if complex cases are also considered. For the remainder of the paper, we considered a value of 90% BS to be the critical value above which incongruence was considered significant.

Taxon Sampling

The number of cases of incongruence supported by more than 90% BS was positively correlated with taxon sampling (Fig. 5). The mean number of incongruent cases at this critical bootstrap value is 0.4 at a level of 20% taxon sampling, whereas it is 2.1 at 80% taxon sampling.

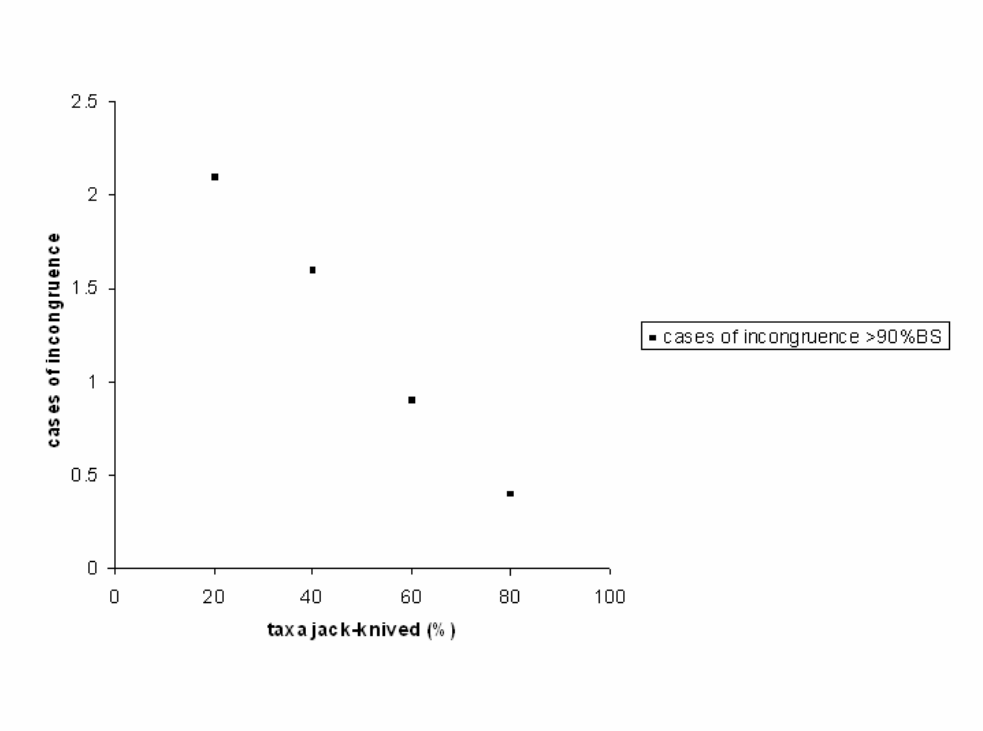


FIG. 5. Relationship between taxon sampling and cases of incongruence that were supported by more than 90% BS.

Long-branch Attraction

The null hypothesis that the difference in number of steps in parsimony analysis of datasets, with and without topological constraints enforced, was caused by the stochastic process of sequence evolution introducing artefacts such as long-branch attraction, was rejected ($P<0.01$) for all taxa (Table 3). The P-value closest to failure to reject the null

hypothesis was found for simulated ITS datasets onto the plastid topology for a sampling including the incongruent *S. monadenum* (P=0.03).

TABLE 3. P-value of the analysis for long-branch attraction with the null distribution generated through parametric bootstrap for simulated plastid and ITS data respectively.

Significantly incongruent accession or clade	PLASTID	ITS
S. parviflorum Eastern Cape	P≤0.01	P≤0.01
Clade 4	P≤0.01	P≤0.01
Clade 3	P≤0.01	P≤0.01
S. monadenum	P≤0.01	P≤0.03
Clade 7	P≤0.01	P≤0.01
Clade 6	P≤0.01	P≤0.01

Orthology/paralogy Conflation, Lineage Sorting, and Hybridization

For orthology/paralogy conflation, the minimum number of ITS duplications and extinctions that needed to be postulated to explain the observed pattern was highest for clade 4 and *S. parviflorum* Eastern Cape (Table 4). For the remainder of the significantly incongruent accessions or clades, one duplication and up to four extinctions sufficed to explain the observed pattern. For lineage sorting, the minimum number of polymorphisms that needed to arise and persist through cladogenic events to explain the observed pattern was also highest for clade 4. For the remainder of the significantly incongruent accessions or clades, only one polymorphism and up to three speciation events sufficed to explain the observed pattern for either, or both, the plastid and ITS datasets (Table 4). For hybridization, apart from clade 6, in each of the remaining seven cases, we needed to postulate dispersal and extinction events (Table 4). In all three cases where the cloned ITS sequences of an accession were not monophyletic, an extinction event of one of the putative parents was needed to explain the observed pattern of multiple copies with hybridization (Table 5). For orthology/paralogy conflation, for *S. rupestre* and *S. fimbriatum* the minimum number of extinct copies of ITS would be one, whereas it would be two for *S. longicauda* var. *longicauda* to explain the observed pattern.

TABLE 4. The minimum number of events that needs to be postulated to explain the observed pattern with orthology/paralogy conflation, lineage sorting, and hybridization respectively.

	orthology/paralogy conflation	orthology/paralogy conflation	lineage sorting	lineage sorting	lineage sorting	lineage sorting	hybridization
	ITS	ITS	ITS	ITS	PLASTID	PLASTID	
significantly incongruent taxon	minimum number of duplications	minimum number of extinction of copies	arisen polymorphisms	speciation events	arisen polymorphisms	speciation events	extinction and dispersal
S. parviflorum Eastern Cape	2	6	2	3	1	3	yes
S. monadenum	1	3	1	2	1	2	yes
Clade 3	1	4	1	3	2	3	yes
Clade 4	2	8	2	5	3	5	yes
Clade 5	1	3	1	2	1	2	yes
Clade 6	1	4	1	3	2	3	no
Clade 7	1	3	1	2	1	2	yes
Clade 8	1	3	1	2	1	2	yes

TABLE 5. The number of events that needs to be postulated to explain the observed topological position of non-monophyletic ITS clones.

cloned species	orthology/paralogy extinction	hybridization extinction/dispersal
S. rupestre	1	yes
S. fimbriatum	1	yes
S. longicauda var. *longicauda*	2	yes

The Species Tree for Satyrium

For the tree resulting from combined plastid and nuclear analysis including all congruent accessions, 85% of all nodes were supported by 50-100% BS, and 79% of all nodes were supported by >95% PP. The topology resulting from the combined analysis was congruent with the topologies resulting from separate analyses of the plastid and ITS dataset. Incongruent accessions or clades were placed in the position that they held in the topologies resulting from separate analysis of the plastid and ITS dataset respectively (Figs.1 and 2). The resulting species tree is generally well-supported and contains several reticulate branches (Fig. 6).

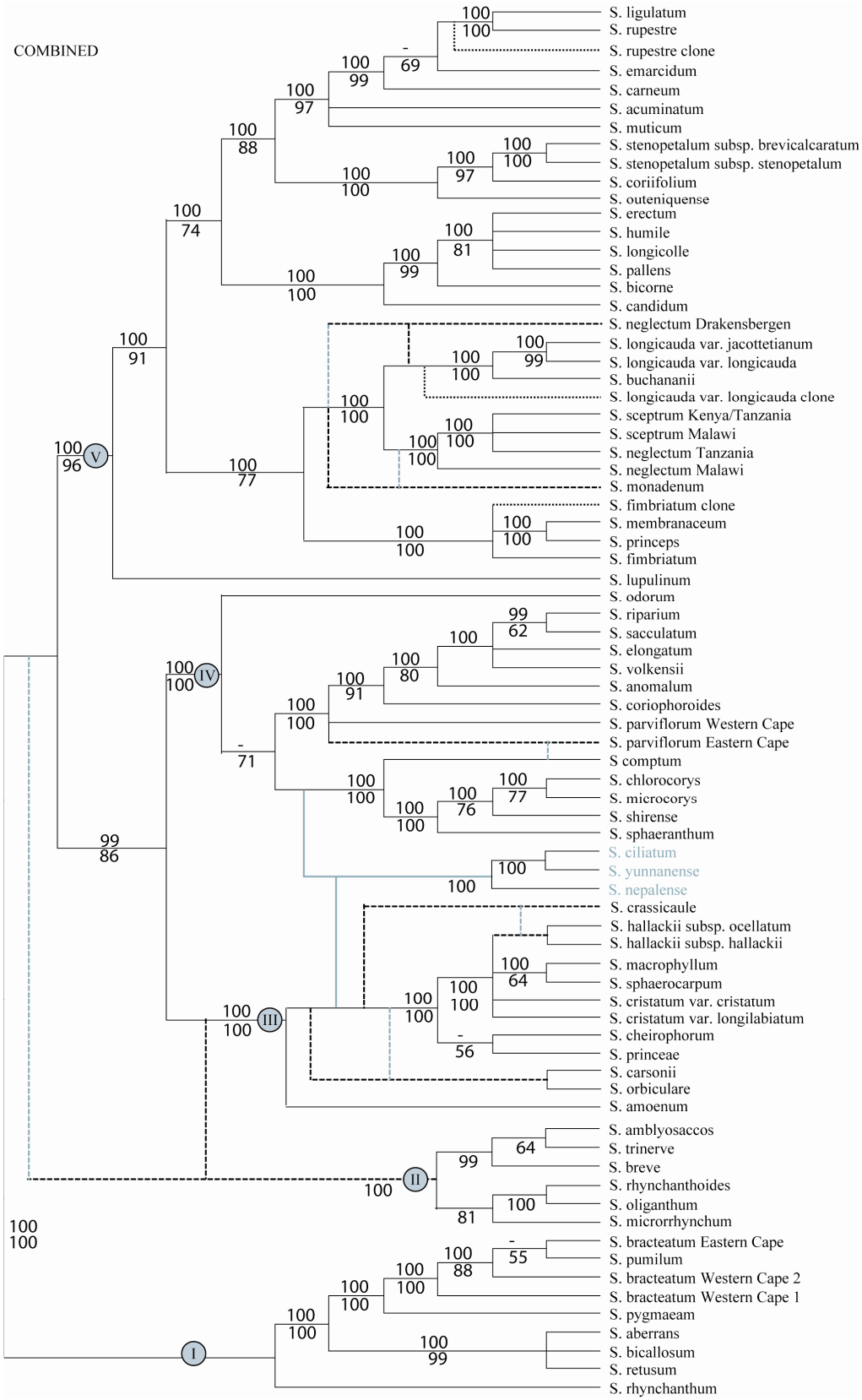


FIG. 6. Species tree for *Satyrium*. Values above branches represent Bayesian PP percentages, values below the branches are bootstrap percentages. Accessions and clades and their subtending branches in grey are putative hybrids that are connected to their respective putative parental branches. Dotted branches represent ITS clones that are connected to their closest relative. Dashed branches represent accessions or clades that cannot be placed with any confidence and for which the cause of incongruence is equivocal. Black dashes connect them to their plastid relatives, grey dashes connect them to their ITS relatives.

DISCUSSION

Separate phylogenetic analysis of the plastid and nuclear dataset of *Satyrium* revealed extensive incongruence. The level of incongruence observed between plastid and nuclear datasets for other angiosperms varies from low (e.g. *Disa* in Orchidaceae; pers. comm. Benny Bytebier), to moderate (e.g. Rubiaceae; Lantz and Bremer 2005), to extensive (e.g. this study; Mason-Gamer and Kellogg 1996). Incongruence between gene trees compromises the straight forward reconstruction of the species tree (Doyle 1992, 1997; Maddison 1997). Here we describe a detailed protocol to deal with incongruence involving three steps: identifying incongruence, assessing the cause of incongruence, and reconstruction of the species tree.

Identifying Incongruence

Assessing significance of incongruent taxa.—Several methods have been proposed to evaluate the significance of incongruence (Templeton 1983; Kishino and Hasegawa 1989; Rodrigo et al. 1993; Farris et al. 1994; Huelsenbeck et al. 1996; Eldenäs and Linder 2000; Swofford 2000; Ronquist and Huelsenbeck 2003). In this study, we have used the ILD test (Farris et al. 1994), implemented as Partition Homogeneity Test in PAUP* 4.0b (Swofford 2000) to assess the significance of incongruence in our datasets. The performance of the ILD test has been criticised (e.g. Yoder et al. 2001; but see Hipp et al. 2004) for a high type I error rate (Cunningham 1997). Several studies have supported this claim with results that show rejection of the null hypothesis of congruence for different datasets that were simulated onto the same topology (e.g. Darlu and Lecointre 2002; Barker and Lutzoni 2002). The question however remains whether the error reported is directly linked to a bias of the ILD test, or whether this is introduced by different biases. The null hypothesis of the ILD test is not parameter-driven and does not make any statements about the process of sequence evolution. All that is tested is whether there is congruence between topologies that are analysed separately using parsimony (Farris et al. 1994). The conditions reported by both Darlu and Lecointre (2002) and Barker and Lutzoni (2002) under which the ILD test was reported to ‘fail’, were those where the parameters of sequence evolution were very heterogeneous between the two datasets that were simulated onto the same topology. These differential patterns of sequence evolution could result in high levels of within-dataset homoplasy. Under these conditions parsimony may fail to retrieve the correct topology due to long-branch attraction, which would ultimately lead to different topologies being estimated for the two datasets under comparison. Although the ILD test is a data-based method, a particular topology is essential in calculating the ILD test P-value. If the above scenario were to apply,

the ILD test would correctly reject congruence, although the underlying phylogenetic history of the data may be the same. Under these circumstances it may then be the failure of the tree reconstruction algorithm to correctly reconstruct the underlying topology, rather than the failed ILD test, which results in the rejection of the null hypothesis of congruence. For this reason we deem the ILD test as suitable to initially detect significant incongruence. Subsequently, the cause of the incongruence can then be assessed, including phenomena such as long-branch attraction, which can result in incongruence despite the underlying phylogeny producing the data, being the same.

Based on the results of the ILD test, we found that several of the accessions or clades that were identified as incongruent in our study did not reach the significance level (P=0.05). These cases were considered to constitute 'soft incongruence' (Seelanan et al. 1997) and therefore more data are needed to test their status.

The relationship between P-values and bootstrap values.—For non-complex cases of incongruence we find that the conflicting phylogenetic position of cases that are not rejected by the ILD test ($\alpha<0.05$), are never supported by more than 86% BS by both datasets. At the same time we find that significant cases are consistently supported by at least 96% BS by both datasets. These values are clearly higher than for instance the 70% BS that is reported in the literature (e.g. Eldenäs and Linder 2000; Goldblatt et al. 2002). Our results would imply that incongruence that is supported by a value of up to at least 85% BS, and possibly as high as 96% BS, by both datasets is still non-significant, as determined by the ILD test. The situation is more complicated with complex cases of incongruence, as bootstrap support values no longer accurately predict the significance of incongruence. We argue that this is due to confounding effects of multiple incongruent accessions or clades. As most studies that apply a bootstrap cut-off value to evaluate incongruence are carried out without prior removal of incongruent accessions or clades, taking a more conservative approach with values below 85% BS is most likely better (e.g. Eldenäs and Linder 2000; Goldblatt et al. 2002). At the same time, if the observation of incongruent accessions or clades leads to removal of these prior to the combined analysis, too low threshold values could result in discarding accessions or clades and, thereby, evidence.

Assessing the Cause of Incongruence

Human error.—There are two reasons why we do not believe that the incongruence observed between our datasets is caused by human error. Firstly, although some incongruent accessions were re-amplified and re-sequenced, the initial result did not change. Secondly, if

we assume that human error would be distributed randomly among datasets, we would expect it not only to occur between the plastid and the ITS datasets but also within the plastid dataset. However, we found no evidence of this.

Taxon sampling.—Wendel and Doyle (1998) suggested that taxon sampling could cause incongruence. We investigated this by counting the number of cases of incongruence at different levels of taxon sampling. Instead of finding a negative correlation between cases of incongruence and taxon sampling, which would indicate that sparse taxon sampling could cause incongruence, we found a strong positive correlation. A simple explanation for this pattern could be that by randomly removing accessions from the dataset, accessions that are incongruent are also removed. The more accessions removed, the greater the chance that incongruent accessions are also removed. Our findings are consistent with those of Mitchell et al. (2000) who found that *increased* taxon sampling generated greater disagreement between two molecular datasets. In contrast, Stockley et al. (2005) found that *decreased* taxon sampling led to more incongruence between datasets comprised of morphological and molecular data. They explain that deletion of crucial accessions leads to homoplasy in some morphological characters, thus resulting in spurious phylogenetic relationships. Comparison of two datasets of a very different nature (i.e. morphological vs. molecular) may be more prone to the effects of taxon sampling on incongruence, than are those of similar datasets (e.g. molecular vs. molecular).

Long-branch attraction.—If long-branch attraction results in erroneous phylogenetic reconstruction for multiple datasets in different ways, incongruence will occur. This could for example be a result of different patterns of sequence evolution between loci or genomes (Wendel and Doyle 1998; Moreira et al. 2002). Several studies have suggested that long-branch attraction caused incongruence. The most convincing cases are those where other causes of incongruence can be excluded *a priori*, such as incongruence between topologies reconstructed from non-recombining loci (e.g. Kennedy et al. 2005), or incongruence between topologies reconstructed from different codon positions of the same locus (e.g. Sanderson et al. 2000). However in addition, long-branch attraction has also been suggested to cause incongruence between different datasets, such as between plastid and nuclear data (Von Hagen and Kadereit 2002). However, even though the pattern of sequence evolution was different between our plastid and nuclear datasets, we suggest that long-branch attraction was not the cause of the observed incongruence in our analyses for two reasons: (1) Maximum-likelihood methods that incorporate explicit modelling of the pattern of sequence evolution, such as Bayesian inference, are thought to be less prone to long-branch attraction than are

parsimony methods (Felsenstein 1978; but see Gaut and Lewis 1995). However, we found no difference between topologies reconstructed using parsimony and Bayesian analyses respectively. (2) The observed difference in number of steps between rival contraint plastid and ITS topologies and MPTs was much higher than the difference in datasets simulated using parametric bootstrapping. Thereby long-branch attraction is clearly refuted as an explanation for the observed incongruence. The generated null distribution used for testing is fully dependent on the particular parameters used for the parametric bootstrap, including the tree topology. We have used a randomly selected, most parsimonious plastid tree to simulate ITS data and vice versa. Thus, in order to test whether our ITS dataset was affected by long-branch attraction we assumed that the plastid tree was the true tree and vice versa. If neither the plastid tree, nor the ITS tree was correct for the particular case of incongruence being tested, then we can not exclude long-branch attraction as explanation for the observed incongruence. However, we regard our tree selection procedure for the parametric bootstrap to be the only logical solution and therefore suitable for this method. Using the same approach, Wiens and Hollingsworth (2000) have shown that long-branch attraction may explain their observed incongruence. Our results suggest that a combination of a heterogeneous pattern of sequence evolution and incongruence should not automatically lead to the conclusion that long-branch attraction is the cause of the incongruence (e.g. von Hagen and Kadereit 2002), as long as it is not tested for explicitly.

The results of the taxon sampling experiment and analyses testing for long-branch attraction suggest that non-biological analytical artefacts are not the cause of the observed incongruence in our dataset.

Orthology/paralogy conflation.—A combination of the presence of paralogous gene copies and gene copy extinction or undersampling can lead to incongruence (Wendel and Doyle 1998). Given the results from our cloning experiments, which revealed the presence of paralogous copies of ITS, there is potential for orthology/paralogy conflation in our dataset. It is reasonable to assume that this would only occur in ITS as very little evidence points towards gene duplication of plastid genes (but see Pirie et al. in prep.).

Our pattern-based method that counts the minimum number of events that needs to be postulated to explain the observed pattern with orthology/paralogy conflation, reveals that for at least one accession and one clade the minimum number of events is quite large. Two duplications and six and eight independent extinctions are required to explain the incongruence for *S. parviflorum* Eastern Cape and clade 4, respectively. We therefore argue that it is unparsimonious to explain their incongruence with orthology/paralogy conflation,

especially when compared to the minimum number of events required by alternative processes (see below). For all other significantly incongruent accessions or clades, one duplication event in combination with up to four independent extinction events are sufficient to explain the observed incongruence. Given the labile nature of the ITS locus (Alvarez and Wendel 2003), we argue that orthology/paralogy conflation cannot be excluded as a process that may have led to the observed incongruence for these taxa. On the same grounds, for all three cases where several clones of an accession were not monophyletic, we also regard orthology/paralogy conflation as the most likely explanation for this, even though at least one extinction needs to be invoked to explain this observed pattern. Alternatively, our assumption that the PCR method should result in identification of all ITS copies present was violated (Buckler et al. 1997; Mason-Gamer 2004).

Part of the reason why orthology/paralogy conflation should be strongly considered as an explanation for our results is because of the nuclear marker used in our study. ITS may be an unsuitable biparentally-inherited marker to choose for reconstruction of a species tree, being plagued with paralogous copies due to its very nature (e.g. Buckler et al. 1997; Hartmann et al. 2001; Ko and Jung 2002). In a critical review, Alvarez and Wendel (2003) questioned the suitability of ITS as a phylogenetic marker citing orthology/paralogy conflation as one of the main reasons. We agree that phylogenetic inference would be severely compromised if ITS were the sole marker used. On the other hand we argue that a combination of markers, including ITS, could reveal additional information for reconstructing the species tree compared to the use of single markers. In addition, ITS is a well known phylogenetic marker that is simple to amplify (Baldwin et al., 1995) and proved to be highly variable in the Orchideae (e.g. Douzery et al. 1999; Van der Niet et al. 2005). Finally, an attempt to use a low copy nuclear marker failed for *Satyrium* probably due to PCR-mediated chimera (Cronn et al. 2002). Therefore, despite its shortcomings, ITS may still be a valuable phylogenetic marker if used in combination with plastid markers.

Lineage sorting.—Unlike orthology/paralogy conflation, lineage sorting is a process that could potentially affect both the plastid as well as the nuclear dataset. However, it is considered to occur less frequently in uniparentally-inherited, non-recombining plastid DNA given its fourfold smaller effective population size (Moore 1995). For lineage sorting to occur, intra-specific polymorphism is required, which was sometimes observed in the few cases where we sampled species from geographically separated areas throughout their range.

By counting the minimum number of events that needs to be postulated to explain the observed pattern with lineage sorting, we obtain a similar result as for orthology/paralogy

conflation. The number of events is highest for clade 4. For all other significantly incongruent taxa the number of events fall within similar ranges of generally one polymorphism and two to three speciation events, for at least the ITS data. As with orthology/paralogy conflation, we therefore regard lineage sorting to be an unlikely explanation for the incongruence of clade 4, but not for the remaining accessions or clades.

For phylogenetic studies carried out above the species level, lineage sorting is often mentioned as a potential cause of incongruence (e.g. McCracken and Sorenson 2005) although this has rarely been established with confidence (e.g. Takahashi et al. 2001). It both requires very thorough taxon sampling as well as character sampling, and these conditions are more easily met in phylogeographic studies (e.g. Olsen and Schaal 1999; Morando et al. 2004). Therefore phylogenetic studies have only limited power in the explicit testing of hypotheses of lineage sorting. Our study exemplifies this by suggesting that it can only be considered unlikely for one out of eight cases of significant incongruence. This uncertainty is very much in line with results from similar studies (e.g. Oh and Potter 2003; Archambault and Bruneau 2004) and cannot be solved unless more data in the form of both taxa as well as characters are collected.

Hybridization.—Hybrid speciation is thought to be a common mechanism in plants (Hegarty and Hiscock 2005). It is therefore relevant to the species tree and should be represented as a reticulate event (Linder and Rieseberg 2004). In *Satyrium*, hybrids are regularly encountered in the field (personal observation; Hall 1982; Ellis and Johnson 1999) with the putative parental species sometimes even arranged in different sections (e.g. *S. bicallosum* x *candidum*, known as *S. guthriei*; Hall 1982). Furthermore, Ellis and Johnson (1999) report that several species are interfertile, and at least the majority of species have the same basic chromosome number x=21 (Hall, 1982). These observations suggest a large potential for hybrid speciation in *Satyrium*.

Especially for clade 4, which represents the sole representatives of *Satyrium* in south east Asia, we consider hybridization likely. Both the alternative processes orthology/paralogy conflation and lineage sorting, require a very high minimum number of events. At the same time hybridization could also be considered costly, given this clade's present distribution. It either requires that (1) both hypothetical ancestral putative parental species went extinct in the progeny's distribution range or (2) that the hybrid progeny dispersed to their current range and went extinct in the parental range. Nonetheless, evidence is accumulating that supports a strong colonizing potential of hybrids (Seehausen 2004). Other lines of evidence further strengthen the argument for the hybrid origin of clade 4. The Asian species express several

morphological features that are confined to either parental clade. Most striking is the entrance of the labellum galea. It is narrow in all members of clade III (numbering following Fig. 6) whereas it is wide in all members of clade IV. Following current taxonomy (Bose and Bhattacharjee 1980; Chen et al. 1999), it is narrow in *S. ciliatum*, wide in *S. yunnanense* and both wide and narrow in *S. nepalense*. An attempt to seek molecular support for the hypothesis of a hybrid origin of clade 4 by extensive cloning of ITS of *S. ciliatum* suggested the absence of a typical clade III ITS copy. Further studies should focus on sequencing of low-copy nuclear genes which could possibly shed more light on a hybrid origin.

For all other incongruent accessions or taxa, hybridization is not favored over the alternative explanations such as orthology/paralogy conflation or lineage sorting because of a lack of data. Most phylogenetic studies that unequivocally establish the hybrid origin of an accession or clade use low-copy nuclear genes (e.g. Smedmark et al. 2003; Mason-Gamer 2004) or detect additive patterns of uniquely fixed markers (e.g. Bateman and Hollingsworth 2004; Gravendeel et al. 2004). These methods either failed (sequencing of low-copy nuclear genes) or are unsuitable for detecting ancient hybridization due to too high levels of change subsequent to the hybridization event. For most other studies that use ITS in combination with plastid genes the hybridization scenario remains a hypothesis that requires further testing (e.g. Lantz and Bremer 2005).

It is widely acknowledged that under most circumstances it is difficult to discriminate among orthology/paralogy conflation, lineage sorting, and hybridization when only a pattern of DNA sequences is available (Wendel and Doyle 1998; Avise 2000). Thus often the biology of the species is used to favour one hypothesis over another. In a study of the coral genus *Acropora*, biological factors are suggested to be the reason why hybridization cannot be excluded (van Oppen et al. 2001). However, this strategy fails to explain why lineage sorting would be a less likely hypothesis. Indeed, studies rarely explicitly attempt to distinguish between lineage sorting and hybridization. A tree-based method applying a molecular clock was suggested by Sang and Zong (2000). A similar approach was followed by Doyle et al (2004). The success of their method relies upon the assumption of no extinction. Thus in the case of hybridization, if any of the putative parental species that gave rise to the initial hybrids had gone extinct, then their method would subsequently falsify hybridization.

Our approach, which relies upon counting the minimum number of events to arrive at a specific observed pattern also suffers from certain shortcomings. It merely provides us with a testable hypothesis. We can only apply the parsimony criterion as a discriminator between the counts. This assumes that the events that are counted are rare, and for ITS this assumption

may be violated (Alvarez and Wendel 2003). There are three additional problems associated with the method. Firstly, for lineage sorting we cannot assume, given our limited sampling of populations within species, that we have found all polymorphisms present. Therefore we fail to count the number of extinct polymorphisms, which means that instead we need another type of event to count as a penalty for the parsimony criterion. In our case we used persistence of the polymorphism through speciation events for this, assuming that its occurrence is rare. This assumption is valid given that the more time that has elapsed, the bigger the chances are of fixation of polymorphic neutral alleles (Avise 2000). Secondly, it is very difficult to penalize hybridization. While the number of nodes separating the incongruent position of an accession or clade is of direct influence on the count of events for both orthology/paralogy conflation and lineage sorting, it is irrelevant for hybridization. The reason for this is that the number of speciation events following ancient hybridization did not necessarily affect the potential to hybridize for the putative hypothetical ancestral parental species that produced the hybrid. Therefore, hybridization can only obtain a maximum score of one event, which is that of dispersal and extinction. A solution could be to consider ploidy levels, morphology, and biological barriers which separate the putative hypothetical ancestral parental species (e.g. van Oppen et al. 2001). Unfortunately too few of these data are available for most of the significantly incongruent taxa. The third problem is that there is no biological logic to compare such different events among one another. It is unclear whether the rise and survival of a polymorphism within a population is equally rare as the duplication of an ITS locus, for example.

The Species Tree

Our species tree is the result of a combination of taking advantage of combined sequence data for taxa that are congruent for both datasets, and subsequent grafting of significantly incongruent taxa onto the species tree into the positions that they held in the separate analyses. The results obtained after the addition of both taxa and characters are in line with Van der Niet et al. (2005). However, the addition of more taxa and characters has increased the number of significantly incongruent taxa compared to those of Van der Niet et al. (2005). In some cases this can be attributed to increasing the number of taxa (e.g. *S. parviflorum* Eastern Cape), in others to increasing the number of characters (e.g. Clade II, Fig. 6). Both are expected. If a group of taxa exhibits some incongruence there is no *a-priori* reason to believe that unsampled taxa will not be incongruent. Likewise, cases that were not significantly incongruence before can be strengthened upon collecting more data. This has

implications for other studies that easily dismiss weakly supported incongruence. Before any evolutionary hypotheses can be tested, any incongruence that is present should be investigated further as it may affect the results of the given test.

Our method for dealing with incongruence differs from both conventional methods that either remove incongruent taxa altogether prior to combined analysis (e.g. Alejandro et al. 2005; Kyndt et al. 2005) or that force incongruent data into a bifurcate framework (e.g. Wang et al., 2004), and network reconstruction (reviewed by Vriesendorp and Bakker 2005) that usually results in reticulate trees, by either inputting the raw sequence data or incongruent trees.

We regard our approach to be the most informative compared to the other methods. By deleting taxa or forcing them into a bifurcating framework one either fails to learn about their status or makes huge assumptions regarding the pattern of evolution. On the other hand, published reticulate trees that are the result of network reconstruction approaches often contain so many reticulate branches, even for small numbers of taxa (e.g. Kennedy et al. 2005), that it becomes increasingly difficult to interpret them in the light of a species tree. In this study we have produced a tree that can be used as a template for testing evolutionary hypotheses.

Humphries (1983) stated that we appear to have reached the limits of cladism when it comes to incorporating hybrids into phylogenetic trees. The availability of molecular data has changed this view dramatically. Phylogenetic incongruence between markers with a different pattern of inheritance can be used to demonstrate hybridization (e.g. Rieseberg et al. 1996). However, given that many other processes can also result in phylogenetic incongruence, it is often treated as a problem rather than as a benefit. We have shown that careful examination of incongruence, in combination with the application of several statistical methods, can result in rejection of certain causes and thereby further resolution of the species tree.

ACKNOWLEDGMENTS

We would like to thank Benny Bytebier, Dirk Bellstedt, Bill Liltved, Hubert Kurzweil, Yi-Bo Luo, and Steve Johnson for kindly providing plant material. The National Geographic society, the Swiss Academy of Sciences, and the Georges andAntoine-Claraz Schenkung provided support for fieldwork. Chloé Galley, Cyril Guibert, Berit Gehrke, Elena Conti, Merran Matthews, and Bastienne Vriesendorp provided useful comments. This work was supported by SNSF grant 31-66594.01, "Adaptive and rapid radiation in the Cape flora".

LITERATURE CITED

Alejandro, G. D., S. G. Razafimandibmbison, and S. Liede-Schumann. 2005. Polyphyly of *Mussaenda* inferred from ITS and trnT-F data and its implications for generic limits in Mussaendeae (Rubiaceae). Am. J. Bot. 92: 544-557.

Alvarez, I., and J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phyl. Evol. 29 :417-434.

Archambault, A., and A. Bruneau. Phylogenetic utility of the LEAFY/FLORICAULA gene in the Caesalpinioideae (Leguminosae): Gene duplication and a novel insertion. Syst. Bot. 29: 609-626.

Avise, J. C. 2000. *Phylogeography: The history and formation of species*. Harvard University Press, Cambridge.

Baldwin, B. G., M. J. Sanderson, J. M. Porter, M. F. Wojchiechowski, C. S. Campbell, and M. J. Donoghue. 1996. The ITS region of nuclear ribosomal DNA - A valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. 82: 247-277.

Barker, F. K., and F. M. Lutzoni. 2002. The utility of the incongruence length difference test. Syst. Biol. 51: 625-637.

Barraclough, T.G., and S. Nee. 2001. Phylogenetics and speciation. Trends Ecol. Evol. 16: 391-399.

Bateman, R. M., and P. M. Hollingsworth. 2004. Morphological and molecular investigation of the parentage and maternity of *Anacamptis x albuferensis* (*A-fragrans* x *A-robusta*), a new hybrid orchid from Mallorca, Spain. Taxon 53: 43-54.

Bose T. K., and S. K. Bhattacharjee. 1980. *Orchids of India.* Naya Prokash, Calcutta.

Buckler, E. S., A. Ippolito, and T. P. Holtsford. 1997. The evolution of ribosomal DNA: divergent and phylogenetic implications. Genetics 145: 821-832.

Chen S., Z. Tsi, and Y. Luo. 1999. *Native orchids of China in colour*. Science Press, Beijing.

Cribb P. J., and S. Thomas. 1997. Orchidaceae. In: *Flora of Ethiopia and Eritrea Vol. 6* eds. Edwards, S., S. Demissew, and I. Hedberg. Addis Ababa and Uppsala.

Cronn, R., M. Cedroni, T. Haselkorn, C. Grover, and J. F. Wendel. 2002. PCR-mediated recombination in amplification products derived from polyploidy cotton. Theor. Appl. Genet. 104: 482-489.

Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? Mol. Biol. Evol. 14: 733-740.

Darlu, P., and G. Lecointre. 2002. When does the incongruence length difference test fail? Mol. Biol. Evol. 19: 432-437.

Douzery, E. J. P., A. M. Pridgeon, P. J. Kores, H. P. Linder, H. Kurzweil, and M. W. Chase. 1999. Molecular phylogenetics of Diseae (Orchidaceae): a contribution from nuclear ribosomal ITS sequences. Am. J. Bot. 86: 887-899.

Doyle, J. J. 1992. Gene trees and species trees – molecular systematics as one-character taxonomy. Syst. Bot. 17: 144-163.

Doyle, J. J. 1997. Trees within trees: genes and species, molecules and morphology. Syst. Biol. 46: 537-553.

Doyle, J. J., J. L. Doyle, J. T. Rauscher, and A. H. D. Brown. 2004. Diploid and polyploidy reticulate evolution throughout the history of the perennial soybeans (Glycine subgenus Glycine). New. Phytol. 161: 121-132.

Eldenäs, P. K., and H. P. Linder. 2000. Congruence and complementarity of mophological and trnL-trnF sequence data and the phylogeny of the African Restionaceae. Syst. Bot. 25: 692-707.

Ellis, A. G., and S. D. Johnson. 1999. Do pollinators determine hybridization patterns in sympatric *Satyrium* (Orchidaceae) species? Plant Syst. Evol. 219: 137-150.

Fan, C. Z., and Q. Y. Xiang. 2003. Phylogenetic analyses of Cornales based on 26S rRNA and combined 26S rDNA-matK-rbcL sequence data. Am. J. Bot. 90: 1357-1372.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult, C. 1994. Testing significance of incongruence. Cladistics 10: 315-319.

Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27: 404-410.

Felsenstein, J. 1985. Confidence limits on phylogeneties: an approach using the bootstrap. Evolution 39: 783-791.

Gaut, B. S., and P. O. Lewis. 1995. Success of maximum-likelihood phylogeny inference in the 4-taxon case. Mol. Biol. Evol. 12: 152-162.

Geerinck D. 1984. Orchidaceae (première partie). In: *Flore d'Afrique Centrale* ed. Bamps, P.. Meise: Jardin botanique national de Belgique.

Goldblatt, P., V. Savolainen, O. Porteous, I. Sostaric, M. Powell, G. Reeves, J. C. Manning, T. G. Barraclough, and M. W. Chase. 2002. Radiation in the Cape flora and the phylogeny of peacock irises *Moraea* (Iridaceae) based on four plastid DNA regions. Mol. Phyl. Evol. 25: 341-360.

Gravendeel, B., M. C. M. Eurlings, C. van den Berg, and P. J. Cribb. 2004. Phylogeny of *Pleione* (Orchidaceae) and parentage analysis of its wild hybrids based on plastid and nuclear ribosomal ITS sequences and morphological data. Syst. Bot. 29: 50-63.

Hall A.V. 1982. A revision of the Southern African species of *Satyrium.* Contributions from the Bolus Herbarium 10: 1-142.

Hamilton, M. B. 1999. Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variation. Mol. Ecol. 8: 521-523.

Hartmann, S., J. D. Nason, and D. Bhattacharya. 2001. Extensive ribosomal DNA genic variation in the columnar cactus Lophocereus. J. Mol. Evol. 53: 124-134.

Hegarty, M. J., and S. J. Hiscock. 2005. Hybrid speciation in plants: new insights from molecular studies. New Phytol. 165: 411-423.

Hipp, A. L., J. C. Hall, and K. J. Sytsma. 2004. Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. Syst. Biol. 53: 81-89.

Huelsenbeck, J. P., D. M. Hillis, and R. Jones, R. 1996. Parametric bootstrapping in molecular phylogenetics: applications and performance. Pp 19-46 in: *Molecular Zoology: advances strategies and protocols* eds. Ferraris, J. D., and S. R. Palumbi. New York: Wiley-Liss, Inc.

Humphries, C. J. 1983. Primary data in hybrid analysis. Volume 2, Chapter 6 in: *Advances in cladistics* eds. Platnick, N. I., and V. A. Funk. New York: Columbia University Press.

Johnson, S.D. 1997. Insect pollination and floral mechanisms in South African species of *Satyrium* (Orchidaceae). Plant Syst. Evol. 204: 195-206.

Johnson, S. D., and K. Steiner. 1999. Pollination. Pp. 34-39 in: *Orchids of Southern Africa* eds. Linder, H. P., and H. Kurzweil. Rotterdam: Balkema.

Kennedy, M., B. R. Holland, R. D. Gray, and H. G. Spencer. 2005. Untangling long branches: identifying conflicting phylogenetic signals using spectral analysis, neighbour-net, and consensus networks. Sys. Biol. 54: 620-633.

Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum-likelihood estimate of the evolutionary tree topologies from DNA-sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29: 170-179.

Ko, K. S., and H. S. Jung. 2002. Three nonorthologous ITS1 types are present in a polypore fungus Trichaptum abietinum. Mol. Phyl. Evol. 23: 112-122.

Kocyan, A., Y.-L. Qiu, P. K. Endress, and E. Conti. 2004. A phylogenetic analysis of Apostasioideae (Orchidaceae) based on ITS, trnL-F and matK sequences. Plant Syst. Evol. 247: 203-213.

Kores, P. J., P. H. Weston, M. Molvray, and M. W. Chase. 2000. Phylogenetic relationships within the Diurideae (Orchidaceae): Inferences from plastid matK DNA sequences.

Pp. 449-456 in: *Monocots: systematics and evolution* eds. Wilson, K. L., and D. A. Morrison. Collingwood: CSIRO.

Kurzweil, H., and H. P. Linder. 1998. A phylogenetic analysis of the genus *Satyrium* (Orchidaceae). *Beiträge zur Biologie der Pflanzen* 71: 101-181.

Kyndt, T., B. van Droogenbroeck, E. Romeijn-Peeters, J. Parcemon Romero-Motochi, X. Scheldeman, P. Goetghebeur, P. van Damme, and G. Gheysen. 2005. Molecular phylogeny and evolution of Caricaceae based on rDNA internal transcribed spacers and chloroplast sequence data. Mol. Phyl. Evol. 37: 442-459.

La Croix I., and P. J. Cribb. 1995. Orchidaceae. In: *Flora Zambesica Vol. 11 part 1* ed. Pope, G. V. London : Flora Zambesica Managing Committee.

Lantz, H., and B. Bremer. 2005. Phylogeny of the complex Vanguerieae (Rubiaceae) genera *Fadogia*, *Rytigynia*, and *Vangueria* with close relatives and a new circumscription of Vangueria. Plant Syst Evol 253: 159-183.

Lewis, P.O. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50: 913-925.

Linder, C. R., and L. H. Rieseberg. 2004. Reconstructing patterns of reticulate evolution in plants. Am. J. Bot. 91: 1700-1708.

Linder, H. P., and H. Kurzweil. 1999. *Orchids of Southern Africa*. Rotterdam: Balkema.

Maddison, W.P. 1997. Gene trees in species trees. Syst. Biol. 46: 523-536.

Mason-Gamer, R. J., and E. A. Kellogg. 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst. Biol. 45: 524-545.

Mason-Gamer, R.J. 2004. Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. Syst. Biol. 53: 25-37.

McCracken, K. G., and M. D. Sorenson. 2005. Is homoplasy or lineage sorting the source of incongruent mtDNA and nuclear gene trees in the stiff-tailed ducks (Nomonyx-Oxyura)? Syst. Biol. 54: 35-55.

Mitchell, A., C. Mitter, and J. C. Regier. 2000. More taxa or more characters revisited: Combining data from nuclear protein-encoding genes for phylogenetic analyses of Noctuoidea (Insecta: Lepidoptera). Syst. Biol. 49: 202-224.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation – mitochondrial-gene trees versus nuclear-gene trees. Evolution 49: 718-726.

Morando, M., L. J. Avila, J. Baker, and J. W. Sites. 2004. Phylogeny and phylogeography of the *Liolaemus darwinii* complex (Squamata: Liolaemidae): evidence for introgression and incomplete lineage sorting. Evolution 58: 842-861.

Moreira, D., S. Kervestin, O. Jean-Jean, and H. Philippe. 2002. Evolution of eukaryotic translation elongation and termination factors: variations of evolutionary rate and genetic code deviations. Mol. Biol. Evol. 19: 189-200.

Nylander, J. A. A., F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53: 47-67.

Oh, S. H., and D. Potter. 2003. Phylogenetic utility of the second intron of LEAFY in Neillia and Stephanandra (Rosaceae) and implications for the origin of Stephanandra. Mol. Phyl. Evol. 29: 203-215.

Olsen, K. M., and B. A. Schaal. 1999. Evidence on the origin of cassava: Phylogeography of *Manihot esculenta*. Proc.Natl. Acad. Sci. U.S.A. 96: 5586-5591.

Polunin O., and A. Stainton. 1984. *Flowers of the Himalaya*. Bombay: Oxford University Press.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Rambaut A., and N. C. Grassly. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Computational and Applied Biosciences 13: 235-238.

Rieseberg, L. H., J. Whitton, and C. R. Linder. 1996. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. Acta Bot. Neerl. 45: 243-262.

Rodrigo, A. G., M. Kellyborges, P. R. Bergquist, and P. L. Bergquist. 1993. A randomization test of the null hypothesis that 2 cladograms are sample estimates of a parametric phylogenetic tree. N. Z. J. Bot. 31: 257-268.

Rokas, A., B. L. Williams, N. King, and S. B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425: 798-804.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Sanderson, M. J., M. F. Wojciechowski, J. M. Hu, T. S. Khan, and S. G. Brady. 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. Mol. Biol. Evol. 17: 782-797.

Sang, T., and Y. Zhong. 2000. Testing hybridization hypotheses based on incongruent gene trees. Syst. Biol. 49: 422-434.

Seehausen, O. 2004. Hybridization and adaptive radiation. Trends Ecol. Evol. 19: 198-207.

Seelanan, T., A. Schnabel, and J. F. Wendel. 1997. Congruence and consensus in the cotton tribe (Malvaceae). Syst. Bot. 22: 259-290.

Simmons, M. P., and H. Ochoterena. 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49: 369-381.

Smedmark, J. E. E., T. Eriksson, R. C. Evans, and C. S. Campbell. 2003. Ancient allopolyploid speciation in geinae (Rosaceae): Evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. Syst. Biol. 52: 374-385.

Soltis, D. E., P. S. Soltis, and J. J. Doyle. 1998. *Molecular systematics of plants II*. Boston: Kluwer Academic Publishers.

Stockley, B., A. B. Smith, T. Littlewood, H. A. Lessios, and J. A. Mackenzie-Dodds. 2005. Phylogenetic relationships of spatangoid sea urchins (Echinoidea): taxon sampling density and congruence between morphological and molecular estimates. Zool. Scr. 34: 447-468.

Summerhayes V. S. 1968a. Orchidaceae (part 1). In: *Flora of Tropical East Africa* eds. Milne-Redhead, E., and R. M. Polhill. London: Crown Agents for Oversea Governments and Adminstrations.

Summerhayes V. S. 1968b. Orchidaceae Vol III Part I. In: *Flora of West Tropical Africa* eds. Hutchinson, J., J. M. Dalziel and F. N. Hepper. London: Crown Agents for Oversea Governments and Adminstrations.

Swofford, D.L. 2000. PAUP*: Phylogenetic Analysis Using Parsimony, version 4.0b. Sunderland, Massachusetts: Sinauer Associates.

Takahashi, K., Y. Terai, M. Nishida, and N. Okada. 2001. Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. Mol. Biol. Evol. 18: 2057-2066.

Templeton, A. R. 1983. Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37: 221-244.

Van der Niet, T., H. P. Linder, B. Bytebier, and D. U. Bellstedt. 2005. Molecular markers reject monophyly of the subgenera of *Satyrium* (Orchidaceae). Syst. Bot. 30: 263-274.

Van der Niet, T., and P. J. Cribb. Submitted. Additions to the Orchid Flora of Tropical East and South-central Africa. Kew Bull.

Van Oppen, M. J. H., B. J. McDonald, B. Willis, and D. J. Miller, D.J. 2001. The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? Mol. Biol. Evol. 18: 1315-1329.

Vanderpoorten, A., A. J. Shaw, and C. J. Cox. 2004. Evolution of multiple paralogous adenosine kinase genes in the moss genus Hygroamblystegium: phylogenetic implications. Mol. Phyl. Evol. 31: 505-516.

Von Hagen, K. B., and J. W. Kadereit. 2002. Phylogeny and flower evolution of the Swertiinae (Gentianaceae-Gentianeae): homoplasy and the principle of variable proportions. Syst. Bot. 27: 548-572.

Vriesendorp, B., and F. T. Bakker. 2005. Reconstructing patterns of reticulate evolution in angiosperms: what can we do? Taxon 54: 593-604.

Wang, Z.-Y., P. W. Fritsch, S. Shi, F. Almeda, B. C. Cruz, and L. M. Kelly. 2004. Phylogeny and infrageneric classification of *Symplocos* (Symplocaceae) inferred from DNA sequence data. Am. J. Bot. 91: 1901-1914.

Wendel, J. F., and J. J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. Pp. 297-349 in: *Molecular systematics of plants II* eds. Soltis, D. E., P. S. Soltis, and J. J. Doyle. Boston: Kluwer Academic Publishers.

Wiens, J. J., and B. D. Hollingsworth. 2000. War of the Iguanas: conflicting molecular and morphological phylogenies and long-branch attraction in Iguanid lizards. Syst. Biol. 49: 143-159.

Won, H., and S. S. Renner. 2003. Horizontal gene transfer from flowering plants to Gnetum. Proc. Natl. Acad. Sci. U.S.A. 100: 10824-10829.

Wood J. R. I. 1997. *A handbood of the Yemen Flora*. Kew: Royal Botanic Gardens.

Yoder, A. D., J. A. Irwin, and B. A. Payseur. 2001. Failure of the ILD to determine data combinability for slow loris phylogeny. Syst. Biol. 50: 408-424.

Young, N. D., and J. Healy. 2003. Gapcoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4: 6.

APPENDIX 1. Voucher information and Genbank accession numbers.

Accession	Locality; Voucher
Disperis villosa SW.	South Africa: Romans river; T49
Corycium sp.	South Africa; T42
Dactylorhiza maculata (L.) Soo	Switzerland: Rossberg, Art-Goldau; T218
Gymnadenia conopsea (L.) R. Br.	Switzerland: Rossberg, Art-Goldau; T219
Habenaria keniensis Summerh.	Kenya: Mt. Elgon; T274
Habenaria petitiana T. Durand & Schinz	Kenya: road Timboroa-Tinderet; T276
Holothrix sp.	South Africa: Walker bay reserve; T7
Platanthera chlorantha Cust. ex Reichb.	Switzerland: Rossberg, Art-Goldau; T218
Schizodium satyriodes (L.) Garay	South Africa: Gifberg; T39
Satyrium aberrans Summerh.	Tanzania: Mbeya Peak; T400
Satyrium acuminatum Lindl.	South Africa: Tradouw Pass; T18b
Satyrium amblyosaccos Schltr.	Malawi: Nyika Plateau; HK2004
Satyrium amoenum A.Rich	Madagascar; Hermans 5401
Satyrium anomalum Schltr.	Malawi: Malosa Plateau; T380
Satyrium anomalum Schltr.	Tanzania: Mbarali, Chimala escarpment; BB2528
Satyrium bicallosum Thunb.	South Africa: Bainskloof pass, Paarl; BB2112
Satyrium bicorne (L.) Thunb.	South Africa: Romans river; T46
Satyrium bracteatum (L.f.) Thunb.	South Africa: Bainskloof pass, Paarl; BB2110
Satyrium bracteatum (L.f.) Thunb.	South Africa: Vlakkeberg; T34
Satyrium bracteatum (L.f.) Thunb.	South Africa: Mount Thomas, Stutterheim; BB2191
Satyrium breve Rolfe	Tanzania: Sao Hill, Iringa; BB2175
Satyrium breve Rolfe	Tanzania: Mafinga; BB2167
Satyrium buchananii Schltr.	Malawi: Nyika Plateau; HK2043
Satyrium candidum Lindl.	South Africa: Silvermine nature reserve, Cape Town; T276A

(Appendix 1 continued)

Satyrium carneum (Dryand.) Sims	South Africa: Walker Bay Reserve; T3
Satyrium carsonii Rolfe	Malawi: Nyika Plateau; T195
Satyrium cheirophorum Rolfe	Malawi: Mulanje Mt.; T199
Satyrium cheirophorum Rolfe	Kenya: Londiani; T271
Satyrium chlorocorys Rchb. f. ex. Rolfe	Malawi: Nyika Plateau; HK1969
Satyrium ciliatum Lindl.	China: Sichuan Province; Luo and Luo 728
Satyrium comptum Summerh.	Tanzania: Mbeya Peak; T401
Satyrium coriifolium Sw.	South Africa: Romans River; T47
Satyrium coriophoroides A.Rich	Ethiopia:
Satyrium crassicaule Rendle	Kenya: Timboroa; T273
Satyrium crassicaule Rendle	Malawi: Nyika Plateau; HK2030
Satyrium cristatum sond. var. *cristatum*	South Africa: Elands Heights, Maclear; BB2297
Satyrium cristatum sond. var. *longilabiatum* A.V.Hall	South Africa: Verloren Vallei Reserve, Belfast; BB2252
Satyrium elongatum Rolfe	Tanzania: Sao Hill; BB2500
Satyrium emarcidum Bolus	South Africa: Walker Bay Reserve; T6
Satyrium emarcidum Bolus	South Africa: Cape Town; T279
Satyrium erectum Sw.	South Africa: Gifberg; T41
Satyrium erectum Sw.	South Africa: Groot Swartberge, Prince Albert; BB2062
Satyrium fimbriatum Summerh.	Kenya: Mount Longonot; T270
Satyrium hallackii subsp. *hallackii* Bolus	South Africa: Betty's Bay, Caledon; Mostert L387
Satyrium hallackii subsp. *ocellatum* (H.Bol.) A.V.Hall	South Africa: Verloren Vallei Reserve, Belfast; BB2258
Satyrium humile Lindl.	South Africa: Bainskloof; T22
Satyrium ligulatum Lindl.	South Africa; HK1815
Satyrium ligulatum Lindl.	South Africa: Red Hill; T278A
Satyrium longicauda Lindl. var. *jacottetianum* (Kraenzl.) A.V. Hall	South Africa: Verloren vallei Reserve, Belfast; BB2249

(Appendix 1 continued)

Satyrium longicauda Lindl. var. *longicauda*	South Africa: Verloren vallei Reserve, Belfast; BB2253
Satyrium longicolle Lindl.	South Africa: Road Robinson pass-Mosselbay; T307
Satyrium lupulinum Lindl.	South Africa: Ookaapseweg, Cape Town; T277A
Satyrium macrophyllum Lindl.	South Africa; Johnson s.n.
Satyrium membranaceum Sw.	South Africa; HK1822
Satyrium microcorys Schltr.	Malawi: Nyika Plateau; HK2015
Satyrium microrrhynchum Schltr.	South Africa: Garden Castle Reserve, Underberg; BB2276
Satyrium monadenum Schltr.	Malawi: Nyika Plateau; T173
Satyrium muticum Lindl.	South Africa: Mossel Bay; WL 802-1
Satyrium neglectum Schltr,	South Africa: Drakensbergen, Mount aux Sources; T164
Satyrium neglectum Schltr.	South Africa: Drakensbergen, Naudesnek; T206
Satyrium neglectum Schltr.	Malawi: Mount Mulanje; T200
Satyrium neglectum Schltr.	Tanzania: Kitulo Plateau; T417
Satyrium nepalense D.Don	Chase O-539
Satyrium odorum Sond.	South Africa: Western Cape Province; T59
Satyrium oliganthum Schltr.	Malawi: Mulanje Mt.; T202
Satyrium orbiculare Rolfe	Malawi: Nyika Plateau; HK2043
Satyrium outeniquense Schltr.	South Africa: Montigu Pass; T306
Satyrium pallens S.D. Johnson & H. Kurzweil	South Africa: Besemfontein nature reserve; T21
Satyrium parviflorum Sw.	South Africa: Eastern Cape, Ntsikeni, Umzimkulu; BB2217
Satyrium parviflorum Sw.	South Africa: Western Cape, Port Elizabeth; T317
Satyrium princeae Kraenzl.	Malawi: Nyika Plateau; HK2005
Satyrium princeps Bolus	South Africa: Cape St. Francis; T357
Satyrium pumilum Thunb.	South Africa: Skoonvlei, Ceres; BB2012
Satyrium pygmaeum Sond.	South Africa: Langeberg; T311

(Appendix 1 continued)

Satyrium retusum Lindl.	South Africa: Swartberg; T296
Satyrium rhynchanthoides Schltr.	Malawi: Nyika Plateau; T186
Satyrium rhynchanthum Bolus	South Africa: Kleinmond nature reserve, Caledon; BB2155
Satyrium riparium Rchb.f.	Malawi: Nyika Plateau; T381
Satyrium rupestre Schltr.	South Africa: Swartberg; DUB 503
Satyrium sacculatum Rolfe	Malawi: Nyika Plateau; T192
Satyrium sceptrum Schltr.	Malawi: Nyika Plateau; HK1985
Satyrium sceptrum Schltr.	Kenya: Londiani; T272
Satyrium sceptrum Schltr.	Tanzania: Mbeya Peak; T403
Satyrium shirense Rolfe	Malawi: Nyika Plateau; HK1968
Satyrium sphaeranthum Schltr.	Malawi: Nyika Plateau; HK1990
Satyrium sphaerocarpum Lindl.	South Africa; Johnson s.n.
Satyrium stenopetalum Lindl. subsp. *brevicalcaratum* (H. Bol) A.V. Hall	South Africa: Gydo Pass; BB2096
Satyrium stenopetalum Lindl. subsp. *stenopetalum*	South Africa: Tradouw Pass; T17
Satyrium trinerve Lindl.	South Africa: Verloren Vallei Reserve, Belfast; BB2255
Satyrium volkensii Schltr.	Tanzania: Mbozi, Mbeya; BB2177
Satyrium yunnanense Rolfe	China: Sichuan Province; Meng Qian-wan & Liu Ming-chun 923 (20)

On Floral Characters, Phylogeny, and Pollinator Shifts in the Twin-Spurred Orchid Genus *Satyrium*

TIMOTHEÜS VAN DER NIET[1], STEVEN D. JOHNSON[2], AND H. PETER LINDER[1]

[1]*Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008, Zurich, Switzerland*

[2]*School of Biological and Conservation Sciences, University of KwaZulu-Natal, P. Bag X01 Scottsville, Pietermaritzburg 3209, South Africa*

Manuscript

Abstract.—*Satyrium* is a southern hemisphere orchid genus that displays a great deal of morphological diversity . For 25 out of 96 taxa, pollinator observation data are available. These data suggest that almost all observed taxa are pollinated by a single functional pollinator class, either bee-, bird-, butterfly-, Fungus Gnat-, beetle-, Carrion Fly-, noctuid-, or hawkmoth-pollination, and are not pollinated by a mixture of pollinator classes. Here we investigated the relationship between floral characters that are deemed important for pollination and pollinators within a phylogenetic framework. The floral characters contained information that allowed for grouping together of taxa pollinated by the same pollinator class. The floral characters also showed significant phylogenetic structure. However, although the degree of similarity of the floral characters among the taxa was negatively correlated with genetic distance, it was significantly greater among taxa that are pollinated by the same pollinator class, than among members of a clade. Thus floral characters reflected pollinator class more than phylogeny. As expected, genetic distance was significantly greater among members of a clade than among taxa that are pollinated by the same pollinator class. We found that certain floral characters appeared to evolve in a correlated fashion with shifts to certain pollinator classes, most notably with bird-pollination. For better estimating the number of pollinator shifts, we inferred pollinator class for another 35 taxa by using and evaluating Non-Metric Multidimensional Scaling, Distance analysis, Classification Trees, and various measurements and comparisons of floral character similarity. We found that the final assignment of a pollinator class to the unobserved taxa most often agreed with the assignment from the Non-Metric Multidimensional Scaling, and least often with the Classification Tree method. Optimization of pollinator classes onto the phylogeny revealed that all pollinator classes have multiple origins. We observed that the number of pollinator shifts on the phylogeny is 50% of the maximum possible number of shifts. Apart from a bias towards shifts from bee- to beetle-pollination, there was no discernable pattern of directionality among the shifts. We found no significant difference in speciation rates among the different pollinator classes. Taxa pollinated by a certain pollinator class appear to be distributed across habitats, with little bias. We found a positive relationship between the number of taxa ocurring in a certain habitat and the number of pollinator classes among those taxa. We did not find any assocation between habitat- and pollinator-shifts. We interpret our results as evidence that the evolution of floral characters is dependent on the phylogeny, but even more so on the pollinators. The number of pollinator shifts is high compared to other studies. This may be related to the specific pollination conditions in southern Africa in which *Satirium* grows.

Keywords: Cape flora, diversification rates, habitat , Orchidaceae, pollinator inference, pollinator shifts, *Satyrium*, southern Africa, floral syndrome

INTRODUCTION

For over one hundred years an intimate relationship between pollinators, floral morphology and angiosperm diversification has been suggested (Darwin 1885; Grant 1949, 1971; Stebbins 1970; Macior 1971; Kiester et al. 1984). Primary evidence for such relationships comes from two sources: (1) studies that demonstrate the role pollinators play in the shaping of the floral phenotype at the species level (Darwin 1885; Johnson 1997a; Johnson and Steiner 1997; Schemske and Bradshaw 1999; Hodges et al. 2002), family level (Grant and Grant 1965), and angiosperm wide (Grant 1949, 1971; Stebbins 1970; Fægri and Van der Pijl 1979; Grant 1994), and (2) A historical link between pollinators and angiosperm diversification was confirmed when it was shown that angiosperm diversification appeared to have proliferated only after the appearance of a large spectrum of pollinators (e.g. Crepet 1983, 1984; but see Gorelick 2001).

Both these lines of evidence however, fail to demonstrate that pollinator shifts have in fact been frequent and, thereby, have facilitated angiosperm diversification. Also, alternative hypotheses explaining the relationship between floral characters and pollinators, including genetic drift (Lande 1976) and pleiotropy, cannot easily be dismissed in the absence of replication. Studying the interaction between pollinators and floral characters in a phylogenetic context could overcome these problems (Barrett et al. 1996; Weller and Sakai 1999). Despite the plethora of species-level phylogenies (e.g. Bakker et al. 2005), surprisingly few rigorous studies are available that combine phylogenetic and pollinator data (e.g. Armbruster 1993; Goldblatt and Manning 1996; Johnson et al. 1998; Kay et al. 2005; Wilson et al. 2006).

This is probably because pollinator data for large groups of species are sparse, as they are difficult to obtain. Observing a certain insect visiting a plant is not equivalent to establishing that this is a pollinator (e.g. Fishbein and Venable 1996; Waser et al. 1996). As an alternative to direct observations, pollinators could be inferred using the pollination syndromes *sensu* Fægri and Van der Pijl (1979) based on floral characters (e.g. Vogel 1954; Goldblatt and Manning 1998; Kay et al. 2005; Goldblatt and Manning 2006). However, it is not clear to what extent these pollination syndromes accurately predict the pollinator of a given species (e.g. Herrera 1996; Ollerton 1998; Ollerton and Watts 2000; but see Hargreaves et al. 2004).

So far, the available evidence from macro-evolutionary studies suggests that if multiple pollinators are present among a group of plant species, there have been multiple pollinator shifts (e.g. Armbruster 1993; Johnson et al. 1998; Kay et al. 2005; Goldblatt and

Manning 2006; Wilson et al. 2006). This evidence supports the hypothesis that plant diversification for some lineages may have been driven by pollinator shifts. What is still largely unknown is the link between morphological change and pollinator shifts (e.g. Johnson and Steiner 1997; Dupont et al. 2004; Fenster et al. 2004), whether there is any directionality in pollinator shifts (e.g. Stebbins 1970; Johnson et al. 1998; Ollerton and Watts 2000), the relationship between pollinators and plant diversification rates (e.g. Dodd et al. 1999), and the relationship between habitat and pollinator shifts (e.g. Johnson 1997a).

Here we address these issues using the twin-spurred orchid genus *Satyrium*, which occurs mainly in the southern hemisphere. For 25 out of 96 taxa, pollinator data are available. These data indicate that almost all taxa are pollinated by a single functional pollinator class (*sensu* Fenster et al. 2004). This makes *Satyrium* an ideal model system to study the interaction between floral characters and pollinator shifts, within a phylogenetic framework. Recently a once clear-cut relationship between floral characters and pollinators has been called into question (Waser 1998), and is known as Ollerton's paradox (Ollerton 1996). This paradox suggests that most plant species are generalist for pollinators, even though their floral characters suggest otherwise (Herrera 1996; Waser 1996). If this were true it would seriously compromise the possibility of establishing a link between floral characters and pollinators. Johnson and Steiner (2000) showed, however, that the notion that most plant species would be generalist, could be biased towards taxa and/or geographic areas. For some families such as for example Iridaceae and Orchidaceae, or geographical areas such as southern Africa, many species may still be characterized by specialized pollination systems (Johnson and Steiner 2003).

The pollinator classes that have been observed for *Satyrium* include birds (Johnson 1996b, Johnson et al. in prep.), bees (Johnson 1997a, 1997b; Johnson et al. in prep.), carrion flies (Johnson 1997b), Fungus Gnats (Garside 1922), butterflies (Johnson 1997b; Johnson et al. in prep.), noctuid moths (Johnson 1997b; Johnson et al. in prep.), hawkmoths (Johnson 1997a; Harder and Johnson 2005; Johnson et al. in prep.), and beetles (Johnson et al. in prep.) (Table 1). Apart from this wide diversity of pollinator classes, which includes almost all syndromes *sensu* Fægri and Van der Pijl (1979), there is a great deal of morphological diversity represented among the taxa (Vogel 1959; Hall 1982; Johnson 1997b; Kurzweil 1996; Kurzweil and Linder 1998). This diversity is mainly expressed in floral characters such as spur length, rostellum shape and size, floral colour, floral scent and period of scent production, flower size, size and orientation of sepals and petals, and the shape of the entrance of the labellum (Figure 1).

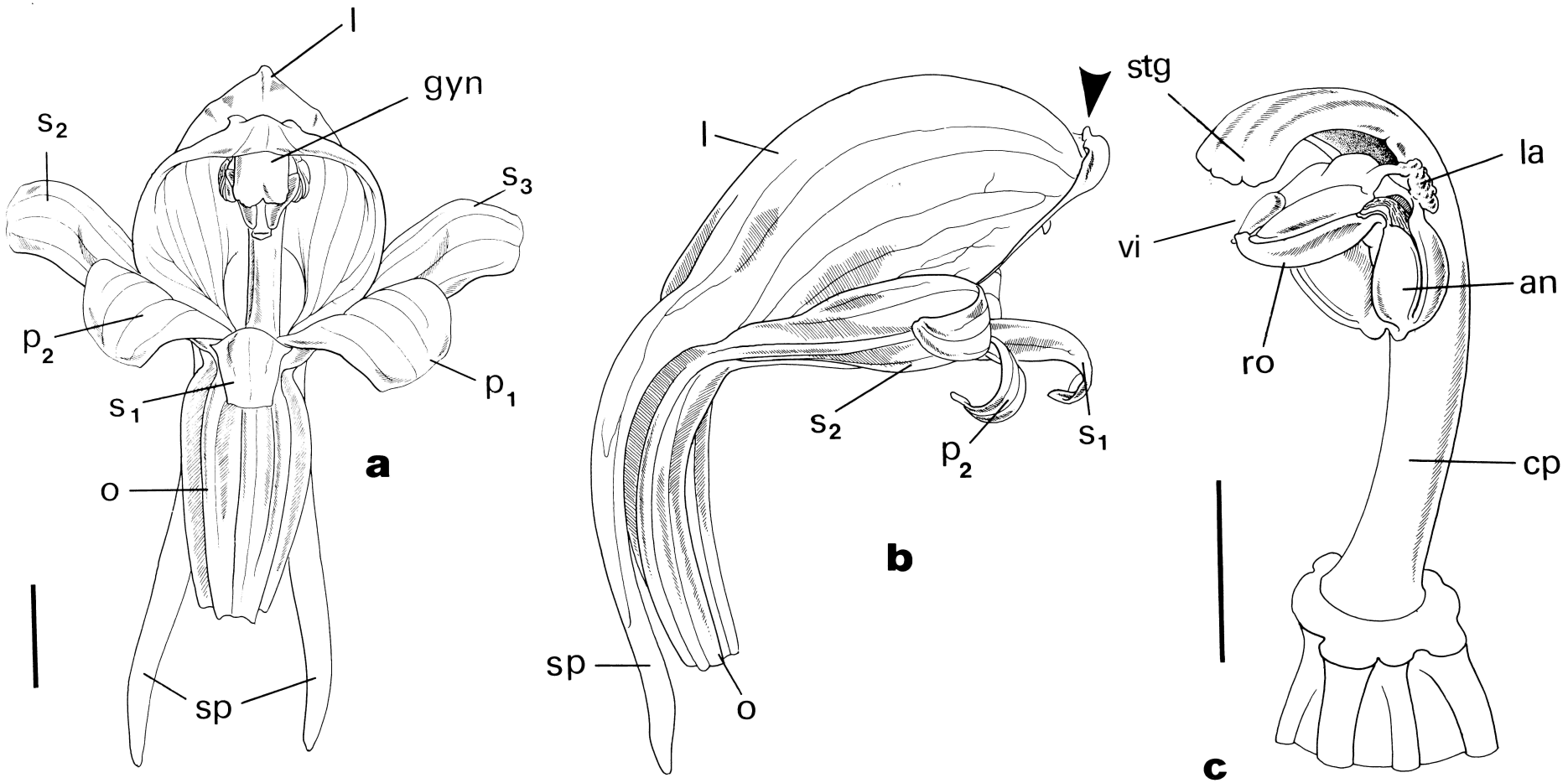


FIG. 1. Floral morphology of *Satyrium carneum*. a-b: flower in ventral and lateral view. c: gynostemium. Abbreviations are as follows: an=anther A_1, cp=column part, gyn=gynostemium, l=lip or labellum, la=lateral apendage of the gynostemium, o=ovary, $p_{1\text{-}2}$=petals, ro=rostellum, s_1=median sepal, $s_{2\text{-}3}$=lateral sepals, sp=spur, stg=stigma flap, vi=viscidium. Reprinted with permission from Kurzweil (1996).

The objective of our study was to investigate the relationship between floral characters, phylogeny and pollinator shifts using a species-level phylogeny (Van der Niet and Linder in review). We address the following questions: (1) to what extent are floral characters related to pollinator classes, as compared to history (phylogeny)? (2) Which, if any, morphological characters evolve in a correlated fashion with shifts to certain pollinator classes? (3) How can we best infer pollinator classes for the taxa for which pollinator observations are lacking? (4) How often do the different pollinator classes originate and are shifts towards a certain pollinator class biased from other pollinator classes? (5) Are diversification rates among the different pollination classes the same? (6) Is there a bias among taxa pollinated by a certain pollinator class to occur in certain habitats and are pollinator shifts associated with habitat shifts?

TABLE 1. Pollinator information for the 25 taxa for which pollinator data are available, and their reference.

Taxon	Pollinator	Reference
Satyrium amblyosaccos	Bee	Johnson et al. in prep.
Satyrium bicallosum	Fungus Gnat	Garside 1922
Satyrium bicorne	Noctuid Moth	Johnson 1997b
Satyrium bracteatum	Carrion Fly	Johnson 1997b
Satyrium carneum	Sunbird	Johnson 1995
Satyrium coriifolium	Sunbird	Johnson 1995
Satyrium crassicaule	Hawkmoth	Johnson et al. in prep.
Satyrium cristatum var. *cristatum*	Bee	Johnson et al. in prep.
Satyrium cristatum var. *longilabiatum*	Bee	Johnson et al. in prep.
Satyrium erectum	Bee	Johnson 1997b
Satyrium hallackii subsp. *hallackii*	Bee	Johnson 1997a
Satyrium hallackii subsp. *ocellatum*	Hawkmoth	Johnson 1997a
Satyrium ligulatum	Noctuid Moth and Butterfly	Johnson 1997b
Satyrium longicauda var. *longicauda*	Hawkmoth	Harder and Johnson 2005
Satyrium membranaceum	Hawkmoth	Johnson et al. in prep.
Satyrium microrrhynchum	Beetle	Johnson et al. in prep.
Satyrium monadenum	Sunbird	Johnson et al. in prep.
Satyrium parviflorum	Noctuid Moth	Johnson et al. in prep.
Satyrium princeae	Butterfly	Johnson et al. in prep.
Satyrium princeps	Sunbird	Johnson 1995
Satyrium rhynchanthum	Bee	K. E. Steiner, pers. com.
Satyrium sceptrum	Sunbird	Johnson et al. in prep.
Satyrium sphaerocarpum	Bee	Johnson et al. in prep.
Satyrium stenopetalum subsp. *brevicalcaratum*	Noctuid Moth	Johnson 1997b
Satyrium trinerve	Beetle	Johnson et al. in prep.

MATERIAL AND METHODS

Taxon Sampling

Taxa that are represented in the phylogenetic species tree of Van der Niet and Linder (in review) were the focus of this study, including all 25 taxa for which pollinator observations are available (observed taxa) (Table 1). Seven taxa were left out of all analyses due a lack of detailed information of their floral characters (*Satyrium amoenum*, *Satyrium ciliatum*, *Satyrium coriophoroides*, *Satyrium elongatum*, *Satyrium emarcidum*, *Satyrium nepalense*, and *Satyrium yunnanense*). One taxon was left out due to poorly understood species-level taxonomy (*Satyrium neglectum*). The final data set consisted of 59 taxa (51 species, 4 varieties, and 4 subspecies) sampled.

Phylogenetic Tree

To investigate the relationship between floral characters and pollinator shifts in a phylogenetic framework, we used the species tree of Van der Niet and Linder (in review) (their Figure 6). This tree contains some uncertainty in the form of reticulations that can be resolved towards the phylogenetic signal in either the plastid or nuclear DNA sequence data. We took this uncertainty into account using both topologies in all analyses that involved a phylogenetic tree. They will be referred to as 'plastid' and 'nuclear' topology, respectively. We included only one accession of *Satyrium sceptrum* and *Satyrium bracteatum* since this would otherwise result in redundancy. To estimate branch lengths of the plastid and nuclear topologies we designed two underlying sequence matrices: the plastid and nuclear matrix. Each matrix included the combined plastid and nuclear DNA sequences. For both the plastid matrix (underlying the plastid topology) and nuclear matrix (underlying the nuclear topology), we deleted the nuclear and plastid sequences respectively, for the taxa that were part of a reticulate branch. In addition, we only included the clonal nuclear sequence that was congruent with the plastid topology if two clones were originally present in the species tree, given the explanation provided by Van der Niet and Linder (in review), that these clones most likely represent gene duplication events instead of hybrid speciation and therefore are irrelevant to the species tree. For both these matrices, including the topologies, we ran Modeltest 3.06 (Posada and Crandall, 1998). The parameters and model provided by the AIC criterion were used to estimate branch lengths and pairwise genetic distances for all taxa for both topologies, using PAUP* 4.0b (Swofford 2000). Ultrametric trees were obtained using Penalized Likelihood and cross-validation of the smoothing parameter (Sanderson 2002) as implemented in the software r8s (Sanderson 2003). Polytomies were resolved according to

secondary knowledge of relationships (Johnson and Kurzweil 1998) or, if that was absent, randomly.

Floral Characters, Pollinator Class, and Phylogeny

Scoring floral characters.—We investigated how floral characters that are putatively involved in plant-pollinator interactions are related to taxa pollinated by the same pollinator class and the phylogeny, respectively. Pollinator classes for these and subsequent analyses were defined according to Fægri and Van der Pijl (1979). We combined the taxa that are pollinated by Fungus Gnats and carrion flies into a category Fly (F). The remaining six classes are Bee (BE), Beetle (B), Bird (BI), Butterfly (BU), Hawkmoth (HA), and Noctuid (N). *Satyrium ligulatum*, which is polymorphic for pollinators, is assigned to the Noctuid class for all analyses.

We scored a set of floral characters that are deemed important for pollination (Johnson 1995; Johnson 1997a; Johnson 1997b; Johnson et al. in prep.) for each taxon. Scoring was based on the investigation of alcohol preserved material and information contained in several monographs (Summerhayes 1968; Hall 1982; la Croix and Cribb 1995; Linder and Kurzweil 1999). All characters were transformed into binary characters for subsequent analyses. If a character was multistate (e.g. colour: red/green/white), each state was considered as a separate binary character (e.g. colour ‘red’: present/absent, colour ‘green’: present/absent, colour ‘white’: present/absent). The final floral character matrix is available from the first author on request.

Floral characters and pollinator class.—We investigated whether floral characters are similar among taxa pollinated by the same pollinator class. First we calculated whether the floral characters contain more information with regard to grouping taxa pollinated by the same pollinator class compared to random groups, for the 25 observed taxa only. We performed Distance analyses of the floral characters in PAUP* 4.0b (Swofford 2000). For all Distance analyses we used the Mean Character Difference to calculate the distance matrix, and the Minimum Evolution (ME) as optimality criterion to reconstruct a tree [default settings in PAUP* 4.0b (Swofford 2000)]. If the floral characters would not contain information that allowed grouping of taxa pollinated by the same pollinator class together, the miminum evolution value would not be smaller among trees where all taxa pollinated by the same pollinator class are constrained to group together, than among random trees. We tested this by randomizing the taxa 100 times, while retaining the structure of the floral character dataset. We carried out both an unconstrained and constrained Distance analysis, applying a backbone

constraint where all the taxa pollinated by the same pollinator class were forced to group together (and also Hawkmoth + Noctuid to group together). We defined the backbone constraint using the shuffled taxon names. We calculated the difference in the ME value between the constrained and unconstrained analysis for each randomization and plotted this in a frequency diagram. The observed difference of the ME value between the constrained and unconstrained analysis of the non-randomized dataset was contrasted to the obtained null distribution in order to test for significance.

Secondly, we performed Non-metric Multi-Dimensional Scaling (MDS) using the floral characters of the observed taxa. All calculations were carried out using NTSYS-pc Version 1.80 (Rohlf 1993). First we calculated the Jaccard similarity of the floral characters. Jaccard similarity is based on the shared presence of a character state and therefore was most suitable for our dataset. We double-centered the Jaccard similarity values. Based on the double-centered values, the first three eigenvectors with their eigenvalues were calculated. The Jaccard similarity and eigenvectors were used as input for the MDS analysis. Plotting of the stress value against the number of dimensions indicated that only three dimensions were sufficient.

Phylogenetic structure of floral characters.—We investigated whether the floral characters contained more structure for inferring phylogenetic relationships compared to random groups. If the floral characters did not contain phylogenetic structure, the Parsimony treelength of trees that are constrained according to the most parsimonious trees derived from phylogenetic analysis of large molecular datasets (Van der Niet and Linder in review) would not be shorter than the treelength of random trees. We tested this by randomizing the taxa 100 times while retaining the structure of the floral character dataset. We calculated the length of the most parsimonious tree in PAUP* 4.0b (Swofford, 2000) in an unconstrained and a constrained parsimony analysis. The topological constraints were according to the most parsimonious plastid and nuclear topology, respectively (Van der Niet and Linder in review). We defined the topological constraint using the shuffled taxon names. The difference in treelength between the constrained and unconstrained analysis was plotted in a frequency diagram for the 100 randomizations. The observed value of the difference in treelength between the constrained and unconstrained analysis for the non-randomized dataset was calculated for the plastid and nuclear topology respectively, and constrasted to the obtained null distributions to test for significance.

Floral character similarity and genetic distance.—We tested whether the evolution of floral characters is phylogenetically labile or conservative. Using the Jaccard similarity of the

floral characters among the observed taxa, we calculated whether this was significantly correlated with genetic distance for both the plastid and nuclear topology. We performed a Mantel test for matrix correlation as implemented in NTSYS-pc Version 1.80 (Rohlf 1993).

We tested whether there is a difference in the Jaccard similarity of the floral characters between the observed taxa pollinated by the same pollinator class and members of the five main clades of Van der Niet et al. (2005), using a t-test. A similar test was also performed for the genetic distance.

Correlated Evolution of Floral Characters and Pollinator Class

We investigated whether pollinators of a pollinator class select for certain floral characters by testing for correlated evolution between floral characters and pollinator shifts. We selected characters from the floral character matrix based on an initial visual assessment. Pollinator class for this analysis is defined as a single pollinator class, or as a group of pollinator classes that are deemed functionally similar for a certain character (e.g. pollinators that receive the viscidia on their body: 'beetle', 'fly', 'bee') (Fenster et al. 2004). Given each selected character for a certain pollinator class, we performed both the Sillén-Tullberg Test (STT) (Sillén-Tullberg 1993) and the Concentrated Changes Test (CCT) (Maddison 1990). For our study, the STT was based on 2x2 contingency table where the rows are branches on a phylogenetic tree that are and are not pollinated by the pollinator class of interest, and the columns represent character stasis (stasis in the ancestral character state on a branch) and character change (change from the ancestral character state to the derived character state on a branch). The null hypothesis of equal patterns of stasis and change, regardless of pollinator class, was tested using the Fisher Exact Test. We used the CCT to test whether gains or losses of a certain character state are more concentrated than can be expected from chance on branches that are optimized for a certain pollinator class. We performed these tests in Macclade 4.07 (Maddison and Maddison 2005) using both the plastid and nuclear topology and ACCTRAN and DELTRAN optimization. If a shift in pollinator class coincided with a shift in a floral character on a certain branch, we always assumed that the pollinator shift preceeded the floral character shift.

Inferring Pollinator Class

To obtain a better estimate of the number of pollinator shifts, we increased taxon sampling by inferring the pollinator class for the taxa for which observations were lacking (unobserved taxa). We performed several analyses and compared the results. For all analyses,

apart from one of the Classification Tree analyses, we used the binary morphological dataset. Given the objective to assign a pollinator class to the unobserved taxa based on floral characters, we weighted the floral characters such that characters containing grouping information were upweighted. Characters that were significantly correlated with certain pollinator classes were candidates for upweighting. We calibrated the weighting scheme using the observed taxa only. By adjusting the character weighting scheme, we minimized the difference in the ME value between an unconstrained and constrained Distance analysis where the taxa pollinated by the same pollinator class were constrained to group together.

Constrained Distance.—We carried out a constrained Distance analysis in PAUP* 4.0b (Swofford, 2000) to place the unobserved taxa among the clusters of the observed taxa. The Distance analysis included all taxa while applying a backbone constraint so that all observed taxa of a certain pollinator class would cluster together. In addition, we constrained Noctuid and Hawkmoth to cluster together. Given only a single purely butterfly-pollinated taxon, we were unable to apply a constraint for the class 'Butterfly'.

Morphological similarity.—We calculated the mean Jaccard similarity of floral characters among taxa pollinated by the same pollinator class for the observed taxa (among-class-similarity). We subsequently calculated the mean similarity of each unobserved taxon to the observed taxa pollinated by the same pollinator class for each pollinator class respectively (to-class-similarity). For each unobserved taxon, we scored whether the to-class-similarity exceeded the among-class-similarity. For these analyses we left the Butterfly class out as the among-class-similarity cannot be calculated for the single butterfly-pollinated taxon.

Non-metric Multidimensional Scaling.— We performed a MDS analysis, as described above, including all taxa. A minimum-spanning tree was calculated based on the Jaccard similarity of floral characters and superimposed on the ordination.

Classification Tree.—To predict the pollinator class for the unobserved taxa, we also used Classification Trees (Breimann et al. 1984). The models were first trained using the data set of the observed taxa. The pollinator class represented the dependent variable and the floral characters were used as a predictor set. In a first analysis we used the raw (non-binary) floral characters for the observed taxa. In a second analysis, we used the three eigenvectors calculated for the MDS as explanatory variables, to asses the effect of compound predictors rather than individual predictive inputs. Model training was performed in the statistical package R using the *tree* library (Ripley 1996; R Development Core Team 2006). The two models were subsequently used to predict pollinators for the unobserved taxa, to classify them and to calculate the proportion of times they were classified into a certain pollinator class.

Final assignment.—We assigned pollinator classes to the unobserved taxa based on a majority-rule criterion, by assigning the pollinator class that was selected most frequently by the individual methods. We also compared the performance of the different methods by calculating the percentage of the final assignment for an unobserved taxon that is congruent with the pollinator class selected by that method. If equal values were obtained in the assignment for an unobserved taxon, we selected the pollinator class that was selected by the method with the highest level of congruence to the final assignment.

Pollinator Shifts

We investigated the number of pollinator shifts and assessed whether there was any directionality among pollinator shifts. We optimized the pollinator classes for the observed taxa only and for all taxa (i.e. including the inferred pollinators) respectively, onto the plastid and nuclear topology using both ACCTRAN and DELTRAN optimization in Macclade 4.07 (Maddison and Maddison 2005). For each pollinator class we calculated from what other pollinator class it has evolved by using the optimizations. We tested whether there was a bias in any of these relationships using a binomial test. We calcluated the predicted proportion of shifts away from a particular pollinator class by summing-up all the internal branches onto which that pollinator class was optimized. This proportion was subsequently contrasted to both the observed number of shifts away from that pollinator class to a pollinator class of interest, and the other shifts towards the pollinator class of interest.

Diversification Rates

We calculated whether there was a significant difference among diversification rates for the different pollinator classes using Kaplan-Meier survival analysis, implemented in SPSS 12.0.1 for Windows. Ultrametric branch lengths were used as a proxy for time to speciation. This is valid since these branch lengths are calculated based on as complete taxon sampling as possible, which equals 70% of the entire genus. Terminal branch lengths were considered as censored data in the survival analyses whereas internal branches were considered as terminating with the event (speciation). Both ACCTRAN and DELTRAN optimizations of pollinator classes were used for the plastid and nuclear topologies. We tested whether there was a significant difference for time to speciation among pollinator classes using the Log Rank test implemented in SPSS 12.0.1 for Windows.

Habitat and Pollinators

To investigate the interaction between pollinator class and habitat, we assigned a habitat class to each taxon (both observed and unobserved) using a large database containing habitat information extracted from labels on herbarium specimens of the K, BR, PRE, and BOL herbaria. We condensed the habitat information down to eight habitat classes: grassland, fynbos, bog, ledge, streamside, coastal, woodland, and bush. To obtain monomorphic coding, each taxon was assigned to one of these eight classes based on a majority rule decision. Using the pollinator class and habitat assignment for each taxon, we plotted the taxon number pollinated by a certain pollinator class per habitat. We also plotted the number of taxa occuring in a certain habitat versus the number of pollinator classes among those taxa. Finally, we optimized habitat in a similar way onto the plastid and nuclear topology, as was done for pollinator classes. We tested whether a pollinator shift was significantly associated with a habitat shift using a Chi-square test.

RESULTS

Floral Characters, Pollinator Class, and Phylogeny

A total of 18 floral characters were selected based on their putative importance for pollination. These were transformed into 45 binary characters (Table 2).

The observed value of the difference in ME value between a constrained and unconstrained Distance analysis of the floral characters, constraining the taxa pollinated by the same pollinator class to group together, is significantly smaller than that of a randomized dataset (Figure 2).

TABLE 2. List of floral characters selected for this study and their states which were transformed into binary characters. The character weight for each character state is indicated.

Character	Binary states	Weight
viscidium size	small	1
	medium	1
	large	3
viscidium shape	globular	2
scent	sweet	2
	putrid	1
	pungent cheese	1
	none	1
scent production	diurnal	1
	nocturnal	1
bract orientation	erect	1

(Table 2 continued)

	spreading	1
	partly reflexed	1
	reflexed	1
labellum colour	white	1
	green	1
	pink	1
	orange	1
	red	1
	yellow	1
galea entrance	narrow	1
	medium	1
	wide	1
rostellum shape	bifid	5
	terminal spreading	1
	lateral	1
sepal- and petal orientation	spreading	1
	reflexed	1
colour contrast	present	2
median sepal length	<10 mm	1
	≥10 mm	1
labellum hairs	present	1
nectar	present	1
oil	present	1
flower nodding	present	2
plant height	<35cm	2
	≥36 cm	2
spur length	0-2 mm	2
	2-12 mm	3
	11-18 mm	3
	>16 mm	3
column length	<2 mm	3
	2-4 mm	3
	5-7 mm	2
	>7 mm	3

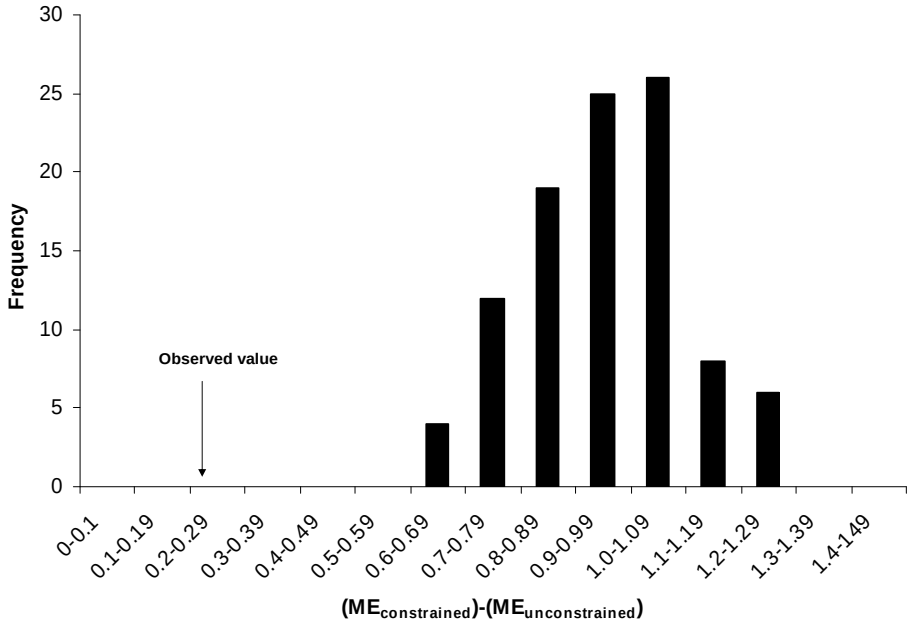


FIG. 2. Frequency of the difference in ME value between a constrained and unconstrained Distance analyses of the floral characters with the observed value plotted.

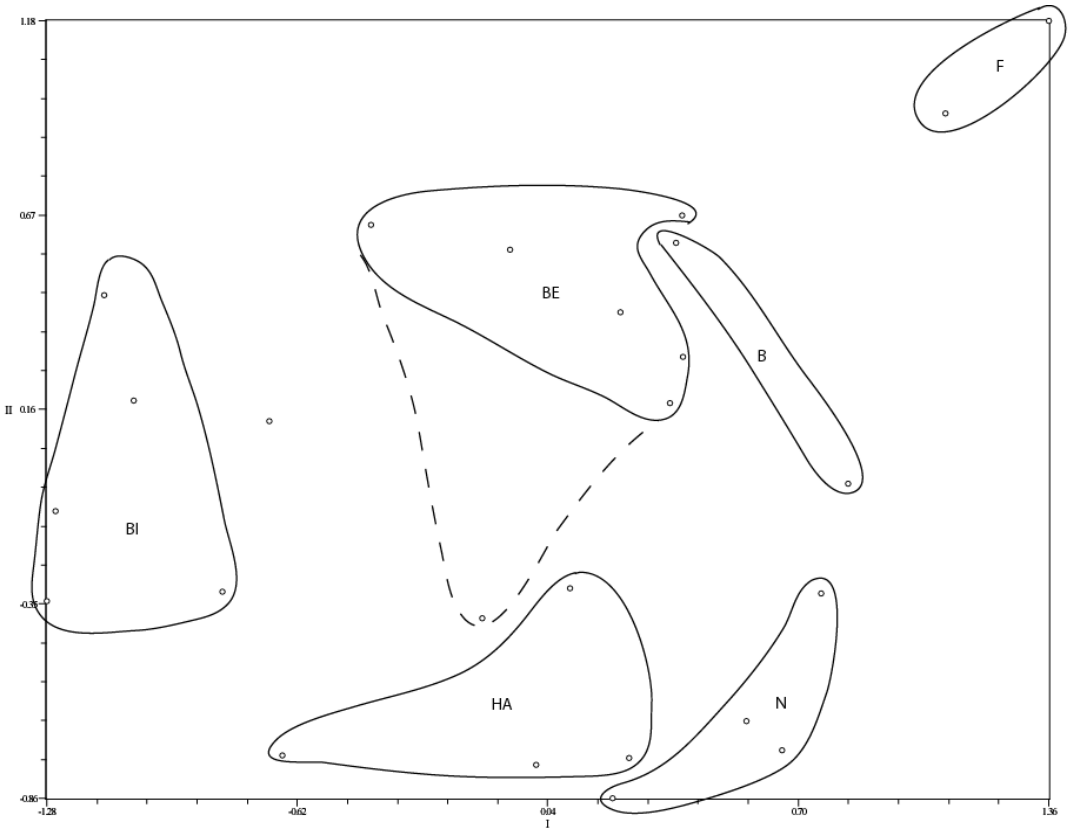
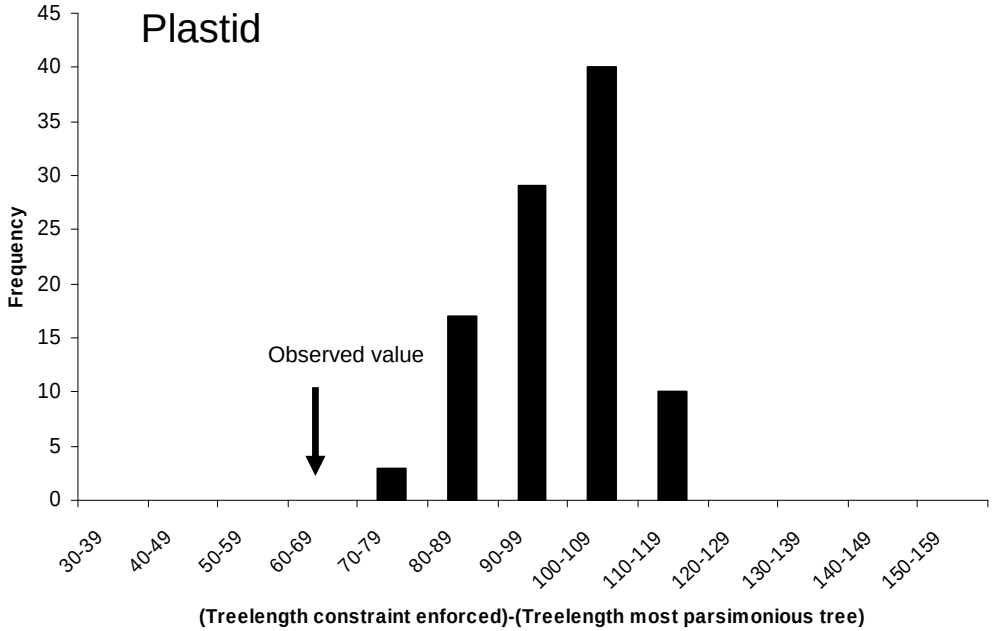


FIG. 3. The first two axes of the MDS analysis of the morphological data for the observed species. Solid lines connect taxa pollinated by the same pollinator classes. The dashed line includes *Satyrium hallackii* ssp. *hallackii* in the Bee class, which is distant from the other members of its class. The only species not part of a class is the sole butterfly-pollinated *Satyrium princeae*.



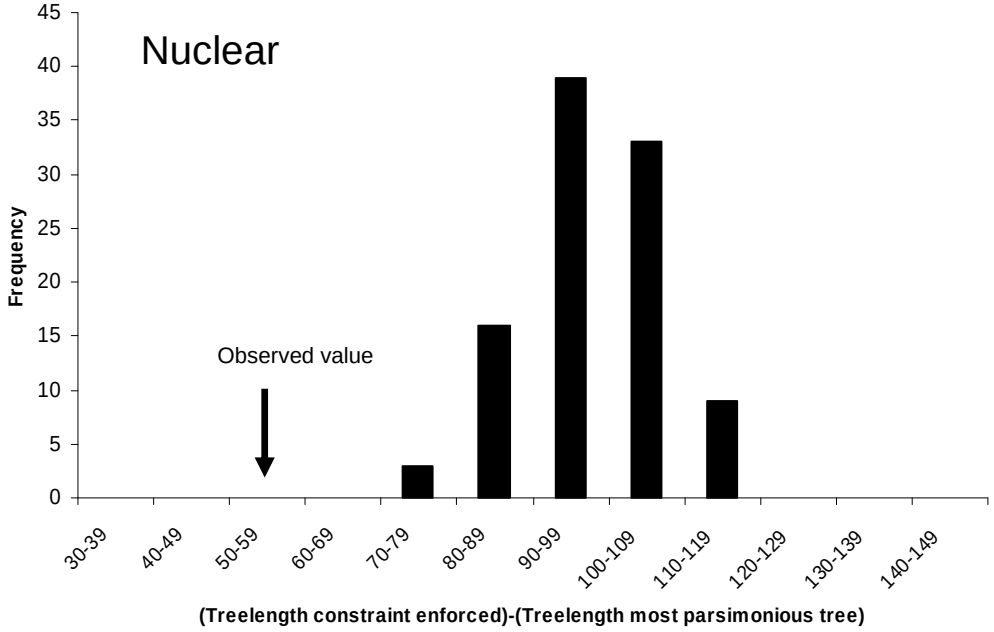


FIG. 4. Frequency diagram for the difference in treelength between random topologies and the most parsimonious tree for the floral characters. The observed value of the difference between a tree constrained according to the plastid and nuclear topology respectively, and the most parsimonious tree is plotted.

The first two axes of the MDS analysis indicate that taxa pollinated by the same pollinator class group together based on the floral characters (Figure 3). The Bird and Fly classes are well separated from the other four classes. There is small overlap between the Bee and Beetle classes and the Hawkmoth and Noctuid classes, respectively. *Satyrium hallackii* ssp. *hallackii*, which is bee-pollinated, groups between the Bee and Hawkmoth class.The floral characters contain more phylogenetic structure than random for both the nuclear and plastid topology (Figure 4). The Mantel test returned a significant result for a (negative) correlation between the plastid genetic distance and Jaccard similarity of floral characters (P=0.002). The result was non-significant for the nuclear genetic distances (P=0.18). Given that the only difference in genetic distances between the plastid and nuclear topology is the result of incongruent taxa, because otherwise the topologies would be identical and therefore genetic distances correlated, we also calculated the correlation between Jaccard similarity of the floral characters and genetic distance on matrices from which the incongruent taxa were removed. This returned a significant result for the plastid genetic distance (P=0.002), as well as for the nuclear genetic distance (P=0.002).

Jaccard similarity of floral characters was significantly larger among taxa pollinated by the same pollinator class than among members of a clade (P=0.00021). At this topological level, the plastid and nuclear topology are congruent and therefore the value for plastid and nuclear topology are the same. Genetic distances were significantly smaller among members of a clade than among taxa pollinated by the same pollinator class for both the plastid topology (P=0.014) and the nuclear topology (P=0.0000118).

Correlated Evolution of Floral Characters and Pollinator Class

Nineteen binary floral characters were selected to test whether their evolution was correlated with a certain pollinator class. Table 3 highlights the characters that were significant for either or both the STT and the CCT. In general the STT returned slightly more significant results than did the CCT. There appears to be no difference in significance between either topology or optimization criterion. Six characters were correlated with bird pollination across topologies, optimization criteria and the statistical test used.

TABLE 3. Correlated evolution between morphological characters and pollinator classes for both topologies and optimization criteria. Only significant (P<0.05) results are indicated (X). If the result was independent of topology or optimization criterion, the **X**'s are highlighted in bold if this applied to either the STT results or the CTT results. **X**'s in bold and underlined represent cases that were significant regardless of topology, optimization criterion or statistical test. All other significant cases are in italics (*X*). '1' means that only one of the two possible tests for each phylogeny or optimization criterion was signficant due to ambiguity in the other test. The categories 'Plastid' and 'Nuclear' summerize the results for both optimization criteria, the categories 'ACCTRAN' and 'DELTRAN' summerize the results for both topologies.

	Sillén-Tullberg Test				Concentrated Changes Test			
Characters and **states** (Pollinator Class)	Plastid	Nuclear	ACCTRAN	DELTRAN	Plastid	Nuclear	ACCTRAN	DELTRAN
colour contrast **present** (BE, BI)		1	1			1	1	
colour **green/white** (N)			*X*				*X*	
colour **pink** (HA, BE)	*X*		*X*					
colour **red/orange** (BI)	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
column length **2-4 mm** (N, HA)	**X**	**1**[1]	**X**[1]	**1**	**X**	**X**[2,3]	**X**[2]	**X**[3]
column length **5-7 mm** (BE)	**X**	**X**	**X**	**X**				
column length **7-10 mm** (BI)	**X**[1]	**X**	**X**[1]	**X**		*X*		*X*
flower nodding **present** (BI)	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
median sepal length **≥10 mm** (BI)	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
rostellum shape **bifid** (BE, B)				*X*				
rostellum size **large** (BI)	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
scent **none** (BI)	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
scent production **nocturnal** (N, HA)	**X**	**X**	**X**	**X**	**X**	**X**	**X**	**X**
spur length **>16 mm** (N, HA)	*X*			*X*				
spur length **11-18 mm** (BI)						*X*		*1*
viscidium shape **globular** (BE, B, CF, FG)	**X**	**X**	**X**	**X**	*X*			*X*

[1]The STT for this character was calculated with alternative column properties: instead of counting the number of branches with stasis of the ancestral state versus branches with a transition from the ancestral to the derived character state, we counted the number of branches where the derived character state was retained versus branches where the derived character state was lost.
[2]In this case we tested whether not *loosing* the character at all on particular branches could be explained by chance alone.
[3]In this case we tested whether the evolution of a *pollinator class* was more concentrated on branches with the given morphological character state. The assumption is that on branches where pollinator shift and character change coincide, the character change preceedes the pollinator shift.

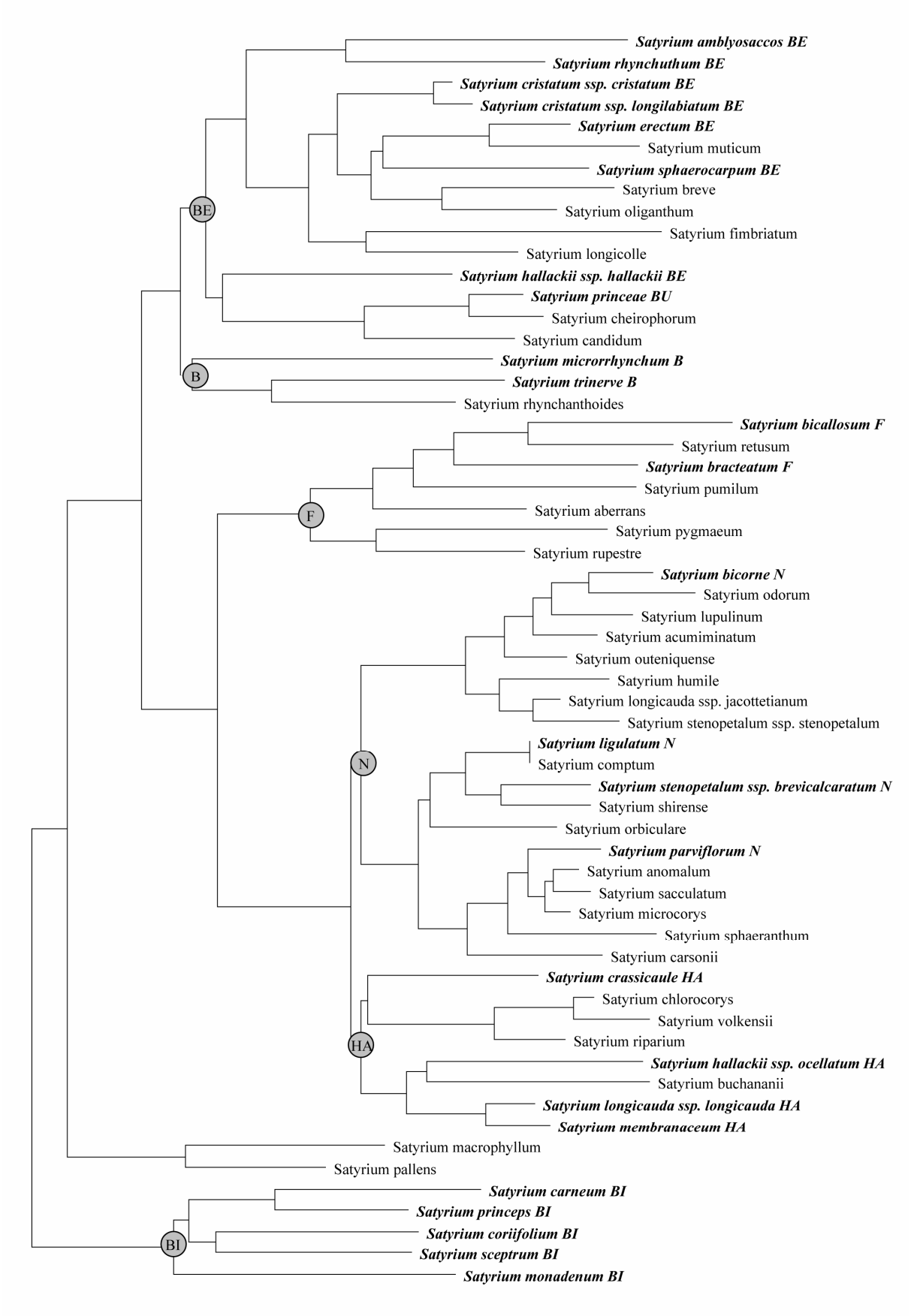


FIG. 5. Clustering of observed and unobserved taxa according to a constrained Distance analysis. Clusters that contain the observed taxa that were constrained to group together are indicated with grey circles. Observed taxa are indicated in bold and italics, with their pollinator classes.

Inferring Pollinator Class

Constrained Distance.—The weighting scheme of the morphological characters is given in Table 2. Results from the constrained Distance analysis show that with the exception of *Satyrium macrophyllum* and *Satyrium pallens*, all taxa fall within the clusters defined by the backbone constraint (Figure 5). Not a single unobserved taxon clustered among the bird pollinated taxa. All other pollinator classes have at least some unobserved taxa nested within them which allowed pollinator assignment for the unobserved taxa (Table 4)

Morphological similarity.—The to-class-similarity for each unobserved taxon shows that in many cases this is highest for the Noctuid class (Table 4). For both the Beetle and Bird class, only one unobserved taxon has the highest to-class-similarity. Few unobserved taxa had a higher to-class-similarity than the among-class-similarity (Table 4). This number was again highest for the Noctuid class. If the Jaccard similarity of floral characters was pooled for Noctuid and Hawkmoth, this resulted in 16 unobserved taxa having a higher to-class-similarity then among-class-similarity for this compound class.

Non-metric Multidimensional Scaling.—Based on the MDS and Minimum-spanning-Tree, all unobserved taxa with the exception of *Satyrium pallens*, were placed in a pollinator class (Table 4, Figure 6). The classes Noctuid + Hawkmoth were treated as a single class for almost all unobserved taxa, unless the MST was unambiguous in distinghuishing between these two classes. If the clustering from the Distance analysis is superimposed on the first two dimensions of the MDS analyses there appeared to be a high level of congruence between the groupings suggested by each analysis (Figure 6).

Classification Tree.—The Classification Trees, using both the raw floral characters data and the first three eigenvectors as predictors, showed that pollinator classes were not always classified together (Figure 7). Using the eigenvectors for instance, there were three terminals for the Bee class. The two models disagreed in about 60% of all cases in their assignment for the unobserved taxa (Table 4).

Final assignment.—For 28 unobserved taxa, the final assignment of a pollinator class was unambiguous based on a majority rule criterion (Table 4). For the remaining six taxa we used the performance of the individual methods to finally assign a pollinator. The majority of unobserved taxa was assigned to the Noctuid class (44%). Not a single unobserved taxon was assigned to the Bird class. Three methods performed approximately equally. The constrained Distance analysis, MDS and minimum-spanning tree analysis, and the method comparing the to-class-similarity to the among-class-similarity all assigned the same pollinator that was selected in the final assignment for about 95% of the cases. The two

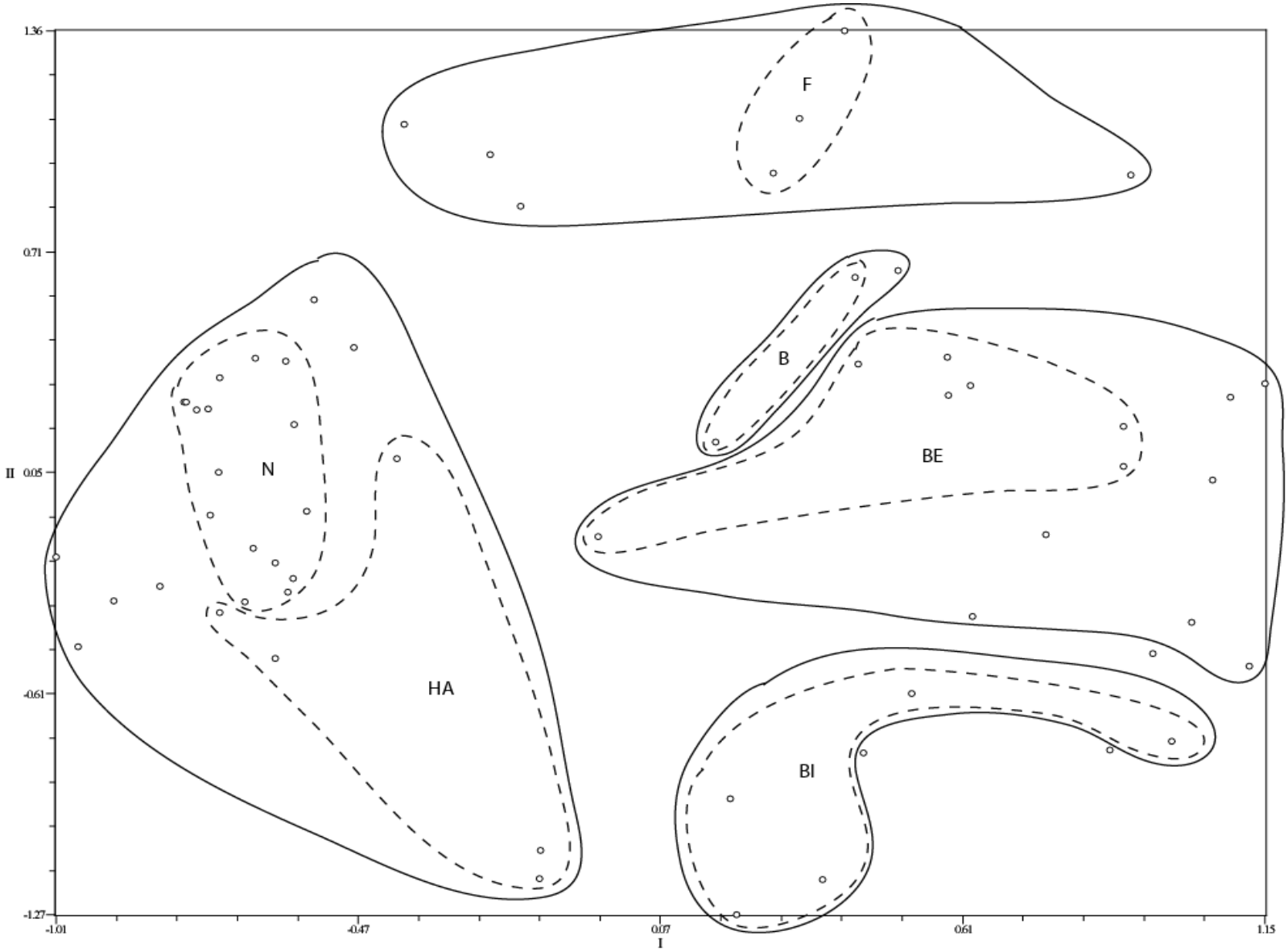


FIG. 6. The first two axes of the MDS analysis based on the floral characters of all taxa. The dashed line represent the observed taxa only. Solid lines represent the clusters from the constrained Distance analysis. Noctuid + Hawkmoth are represented as one cluster.

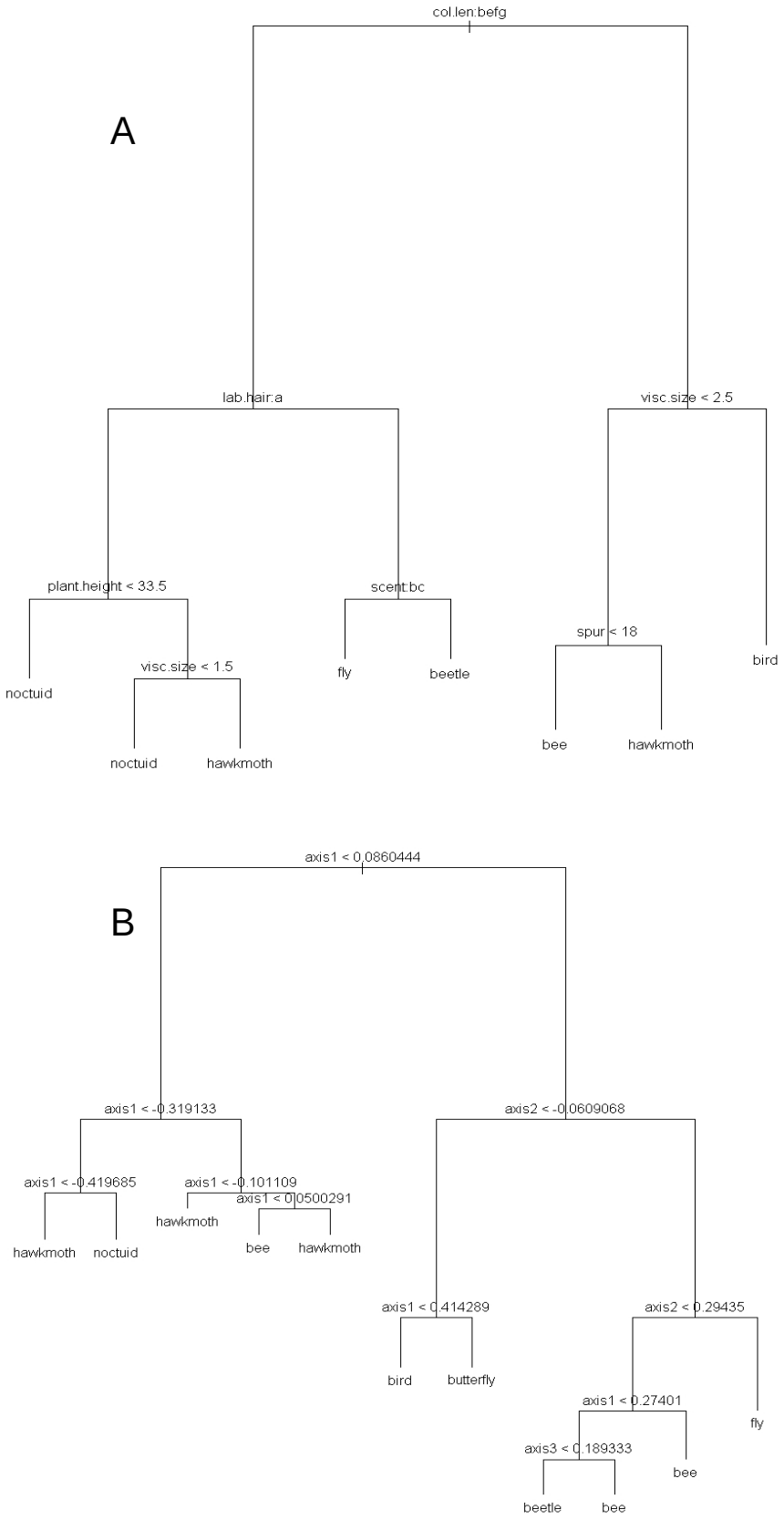


FIG. 7. Classification Tree using floral characters of the observed taxa as predictor values. Decision criteria based on the floral characters are given above the branches. a: Raw data used as input. b: The first three eigenvectors used as input.

TABLE 4. Results from the different methods to infer pollinators for the unobserved taxa. For the taxa for which the final assignment was ambiguous, the pollinator class in bold was assigned based on the performance of the individual methods. The percentage congruent assignments for each individual method compared to the final assignment is given in the final row.

Unobserved taxa	CD[1]	HMS[2]	ACS[3]	MDS+MST[4]	classification RD[5]	classification E[6]	Final
Satyrium aberrans	F	F	-	F	BE/BI	BE	F
Satyrium acuminatum	N	N	N+HA	N+HA	N	N	N
Satyrium anomalum	N	N	N+HA	N+HA	HA	N	N
Satyrium breve	BE	BE	BE	BE	BE/BI	BE	BE
Satyrium buchananii	HA	HA	-	N+HA	BE/BI	BE	HA
Satyrium candidum	BU	BE	-	BE/BU	BE	BI	BE
Satyrium carsonii	N	N	-	N+HA	F	HA	N
Satyrium cheirophorum	BU	BE	-	BI/BU	BE	BU	BU
Satyrium chlorocorys	HA	N	N+HA	N+HA	N	HA	N/**HA**
Satyrium comptum	N	N	N+HA	N+HA	N	HA	N
Satyrium fimbriatum	BE	BE	-	BE	BI	BU	BE
Satyrium humile	N	N	-	N+HA	B	HA	N
Satyrium longicauda var. *jacottetianum*	N	N	N+HA	N+HA	HA	N	N
Satyrium longicolle	BE	BE	BE	BE	BE/BI	BU	BE
Satyrium lupulinum	N	N	N+HA	N+HA	N	N	N
Satyrium macrophyllum	UNPLACED	BI	-	HA/BU	HA	BU	HA/**BU**
Satyrium microcorys	N	N	N+HA	N+HA	HA	HA	**N**/HA
Satyrium muticum	BE	BE	B	BE	BE	BE	BE
Satyrium odorum	N	N	N+HA	N+HA	N	N	N
Satyrium oliganthum	BE	BE	-	BE	BE/BI	BE	BE
Satyrium orbiculare	N	N	N+HA	N+HA	N	HA	N
Satyrium outeniquense	N	N	N+HA	N+HA	N	N	N
Satyrium pallens	UNPLACED	BE	-	UNPLACED	HA	BI	**BE**/BI/HA

(Table 4 continued)

Satyrium pumilum	F	F	-	F/BE	BE/BI	F	F
Satyrium pygmaeum	F	N	-	F	BE/BI	BE	**F**/BE
Satyrium retusum	F	F	F	F	BE/BI	F	F
Satyrium rhynchanthoides	B	B	B	BE	N	F	B
Satyrium riparium	HA	HA	N+HA	N+HA	HA	HA	HA
Satyrium rupestre	F	N	-	N+HA/F	BE/BI	BE	BE/N/**F**
Satyrium sacculatum	N	N	N+HA	N+HA	HA	N	N
Satyrium shirense	N	N	N+HA	N+HA	N	N	N
Satyrium sphaeranthum	N	N	N+HA	N+HA	HA	N	N
Satyrium stenopetalum ssp. *stenopetalum*	N	N	N+HA	N+HA	N	HA	N
Satyrium volkensii	HA	N	N+HA	N+HA	N	HA	HA
Percentage congruent with final	94	88	95	94	59	68	100

[1]constrained Distance analysis
[2]pollinator class for which the to-class-similarity was highest for the unobserved taxa
[3]pollinator class for which the to-class-similarity was higher than the among-class-simililarity.
[4]Non-metric multidimensional scaling and minimum-length spanning tree.
[5]Classification tree using the raw morphological data as predictor
[6]Classification tree using the first three eigenvectors as predictor

methods using a Classifaction Tree only assigned the pollinator that was finally assigned in about 60% of the cases.

Pollinator Shifts

If both the observed and unobserved taxa are considered, all pollination classes had multiple origins, regardless of topology or optimization (Table 5, Figure 8 and 9). Considering only the observed taxa, all pollination classes also had multiple origins with the exception of fly-pollination and beetle-pollination (single origin under ACCTRAN only; results not shown). The origin of beetle-pollination is significantly biased to originate from bee-pollinated lineages (Table 7; binomial test P=0.0324 and P=0.0049 for ACCTRAN and DELTRAN optimization respectively). Between 25 and 30 pollinator shifts, depending on topology and optimization criterion, have taken place on a total of 118 branches. Most shifts were from noctuid-pollination (between 9 and 12 times for both topologies and optimization criteria). Bee-pollination arose most times (between 6 and 8 times for both topologies and optimization criteria).

TABLE 5. Number of pollinator shifts for the plastid topology and DELTRAN optimization. The directionality of change is from the pollinator in rows to the pollinator in the columns. Results for other topologies or optimization criteria were similar.

	BE	B	BI	BU	F	HA	N	SUM
BE	0	2	0	2	0	2	1	7
B	1	0	0	0	0	0	0	1
BI	1	0	0	0	0	2	0	3
BU	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0
HA	0	0	0	0	0	0	2	2
N	4	0	3	0	1	2	0	10
equivocal	1	0	0	0	1	0	0	2
SUM	7	2	3	2	2	6	3	25

Diversification Rates

There was no significant difference of diversification rates among different pollinator classes using survival analysis, regardless of topology or optimization. The Beetle class had the greatest average branch length across topologies and optimization criteria (Figure 10).

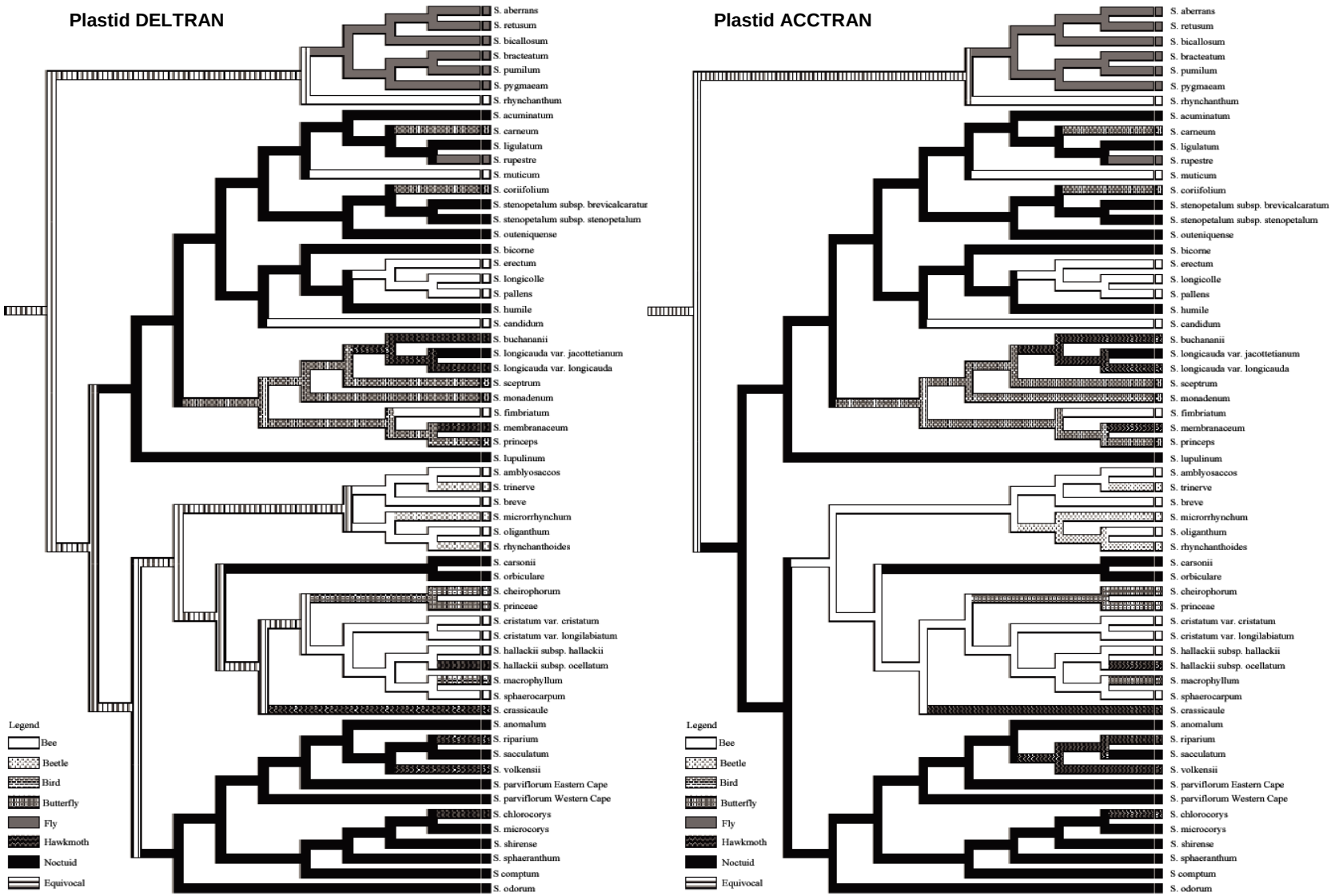


FIG. 8. Optimizations of pollinator classes onto the plastid topology under both DELTRAN and ACCTRAN parsimony optimization.

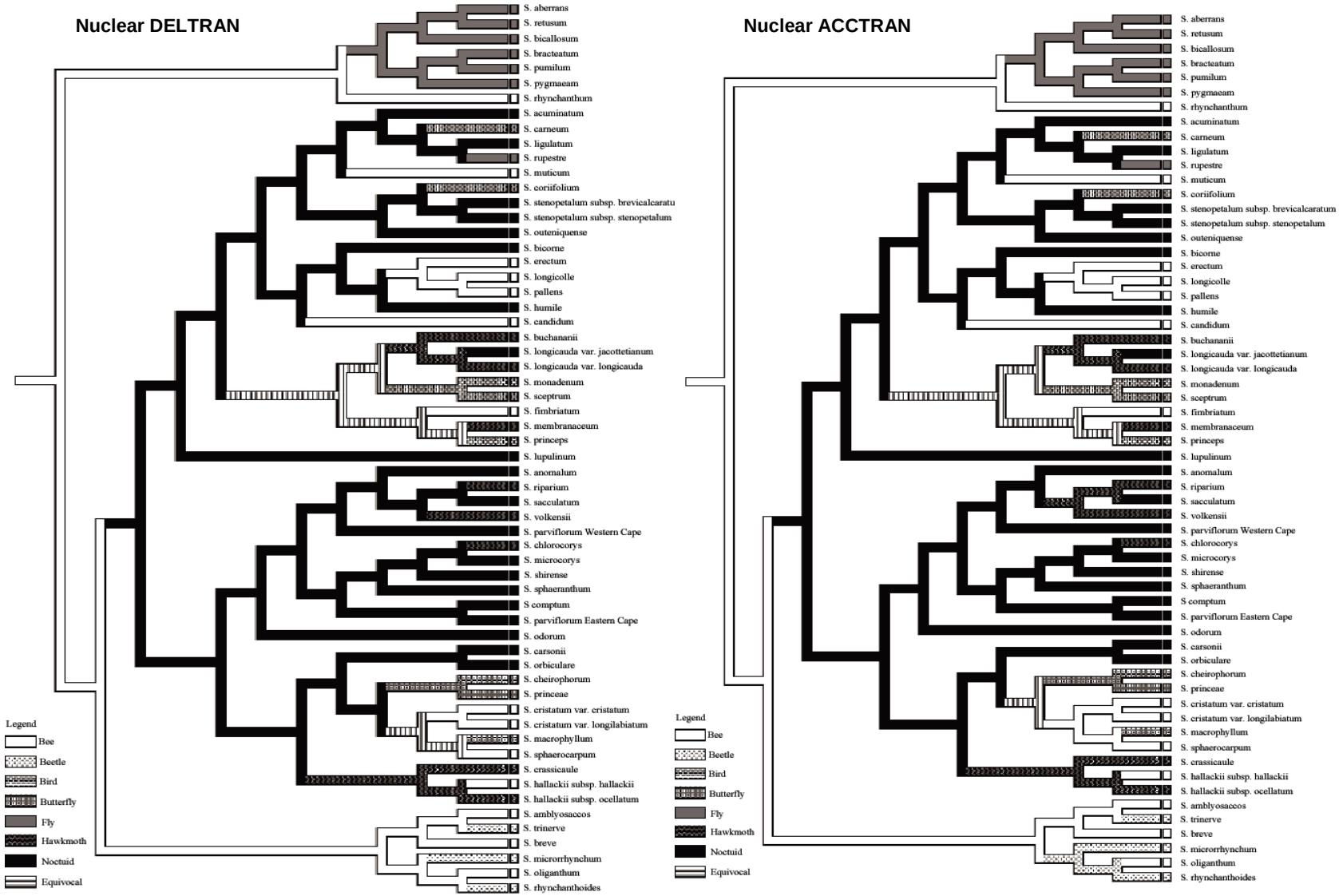


FIG. 9. Optimizations of pollinator classes onto the nuclear topology under both DELTRAN and ACCTRAN parsimony optimization.

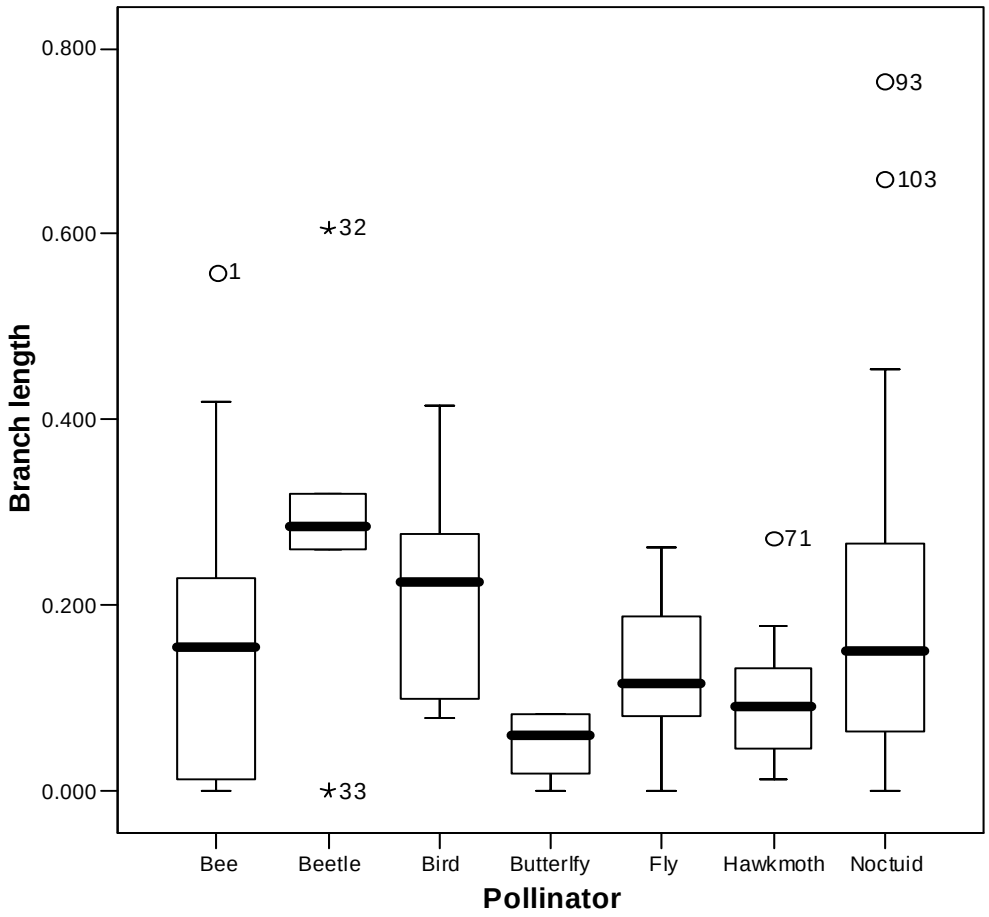


FIG 10. Average branch length with standard deviation for each pollinator class plotted for the plastid topology under ACCTRAN optimization of pollinator classes. Results for other topologies or optimization criteria were similar.

Habitat and Pollinators

Species pollinated by certain pollinator classes were to some extent equally distributed among habitats. The grassland habitat contained taxa that are pollinated by almost all pollinator classes in this study. Moreover, these are more or less represented by equal numbers of taxa. There were few biases, however. Taxa pollinated by Hawkmoths occur only in grassland or bogs. The coastal, bush, and woodland habitats have very few taxa occuring in them and these are all pollinated by a single pollinator class. Fly-pollination is almost confined to taxa occuring on ledges and along streamsides. There is a positive correlation between the number of taxa occuring in a habitat and the number of pollinator classes represented among those taxa (Figure 11). There is no significant association between pollinator shifts and habitat shifts, regardless of topology or optimization criterion used.

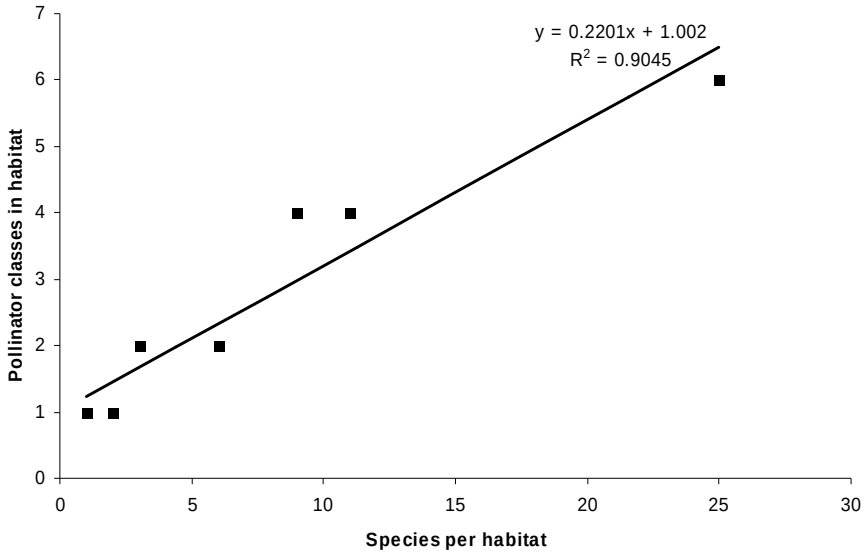


FIG. 11. Correlation between the number of taxa per habitat and number of pollinator classes among those taxa.

Floral Characters, Phylogeny, and Pollinator Class

We have investigated the interactions between floral characters, phylogeny and pollinator shifts for 59 taxa of the orchid genus *Satyrium* which mainly occurs in the southern hemisphere. Our first finding is that floral characters are similar among taxa that are pollinated by the same pollinator class, suggesting that floral syndromes can be defined. This is in agreement with the results of other studies (e.g. Van der Pijl and Dodson 1966; Sakai et al. 1999; Fenster et al. 2004; Hargreaves et al. 2004; Wilson et al. 2004). The existence of floral syndromes was recently called into question (e.g. Herrera 1996; Ollerton 1996). However, this may apply mostly to taxa that are pollinated by several pollinator classes. The pollinator observations for *Satyrium* indicate that almost all taxa are exclusively pollinated by one functional pollinator class (e.g. Johnson 1997b). If taxa attract only one pollinator class, it is likely due to the fact that they possess the floral characters which attract that particular pollinator class but not others. Even if multiple visitors would be attracted, the complicated morphology of the reproductive organs of *Satyrium* would probably only allow for succesful pollination by a pollinator class, which closely matches this morphology (e.g. Vogel 1959; Johnson 1997b). Therefore, given a pollinator's abilities to respond to specific cues and its specific 'bauplan', the existence of floral syndromes could almost be regarded as an obvious consequence of specialization. A case where high pollinator specificity and plant morphology is uncoupled, would be one where the distribution of pollinators is divided so finely within the landscape, that in each habitat only one pollinator resides. We know this is not the case in *Satyrium*, given the many sympatric species that are pollinated by different pollinators. On Nyika Plateau in Malawi, at least seven sympatric taxa that were probably pollinated by five different pollinator classes were encountered.

There is considerable overlap in the grouping of pollinator classes between our results and the angiosperm-wide study of Ollerton and Watts (2000). They also found that both Noctuid (their 'moth') + Hawkmoth, and Bee + Butterfly + Bird group together, respectively. The biggest difference between our study and theirs, is that Beetle is distant from Bee in their analysis, while it overlaps somewhat in our results. The characters included in our study and those of Ollerton and Watts (2000) only overlap to a minor extent. Many characters in our study are specific to the complicated morphology of the *Satyrium* flowers, while Ollerton and Watts (2000) tried to include angiosperm-wide characters. This suggests that, regardless of the characters, the groupings could be considered as natural. Indeed one could argue, for example, that bees and birds are attracted to more similar flowers than are birds and noctuids.

Similarity of floral characters was greater among taxa pollinated by the same pollinator class than among members of a phylogenetic clade. This could have implications for taxonomy and systematics which is mainly based on floral characters. If similarity would be used as a criterion to classify taxa, this could result in classifications that are more congruent with pollinator classes than with evolutionary history. This may have been the case for the neotropical *Costus* where all bird pollinated species were classified into one group. Molecular data have since revealed the polyphyletic nature of this group (Kay et al. 2005). Hall (1982) also grouped the South African *Satyrium* taxa based on morphological similarity. Indeed, some of his clusters conform to pollinator classes, even though he included additional characters and not only those that are putatively involved in pollination. Similar cases are found among the Iridaceae, although Goldblatt and Manning (2006) argue that careful examination of the entire morphology often leads to taxonomy that is congruent with evolutionary history.

The interaction between floral characters, phylogeny, and pollinator class is complex. The floral characters contain both significant phylogenetic structure, as well as significant grouping information with regard to taxa pollinated by the same pollinator class. If pollinator shifts were rare and phylogenetically constrained, this would be obvious. However, this is not the case for *Satyrium*, which is illustrated by the genetic distance, which is significantly smaller among members of phylogenetic clades than it is among taxa pollinated by the same pollinator class. On the other hand, some conservatism in the evolution of floral syndrome characters is suggested. The Mantel test shows that genetic distance, a proxy for phylogenetic distance, is significantly negatively correlated with similarity of floral characters, at least for the plastid topology. If the floral characters were to evolve free from any constraint, this correlation would not be expected. We conclude that the evolution of floral characters is influenced both by pollinator class and by phylogeny. This is supported by the results from the MDS analysis. Although in general, taxa were grouped according to their pollinator class, there were several exceptions. Both *Satyrium hallackii* ssp. *hallackii* and *Satyrium trinerve* group closer to their closest phylogenetic relatives that are pollinated by a different pollinator class, than to taxa that are pollinated by the same pollinator class. The fact that their phylogenetic relatives are pollinated by a different pollinator class suggests that the pollinator shift was relatively recent. Therefore the floral characters may not yet have been shaped optimally for their pollinator class. An alternative explanation to account for the finding that the evolution of floral characters is influenced by both pollinator class and phylogeny, could be that this result was based on examining a suite of floral characters and therefore some

specific characters may correlate with the phylogeny, while others correlate with pollinator class.

Correlated Evolution of Floral Characters and Pollinator Class

We found that the evolution of several floral characters correlated with shifts to certain pollinator classes. However, not all floral characters returned a significant result when placed in a phylogenetic context. On closer examination, apart from colour and scent, characters that varied in size correlated particularly well with shifts to certain pollinator classes. Armbruster (1996) also demonstrated multiple origins of a particular size of an organ (resin gland) that allows for a better fit between the plant and its pollinator, plus a correlation with a shift towards certain pollinators. In *Satyrium,* these characters are column length and spur length. Both the elongate column and the possession of two labellum spurs are synapomorphies for the genus *Satyrium* (Linder and Kurzweil 1994). These characters could therefore be regarded as a key innovation which allowed *Satyrium* to adapt to an array of different pollinators (e.g. Johnson 1997b), analogous to the nectar spurs of *Aquilegia* (Hodges 1997). Interestingly, it appears that it is the *combination* of having labellum spurs and an elongate column that has resulted in the adaptive radiation of *Satyrium*. Two other genera with their center of diversity in southern Africa, that either have an elongate column (*Pachites*) or two spurs (*Disperis*), are not characterized by the presence of many pollination systems (Manning and Linder 1992).

The Bird class has the most floral characters that have evolved in a correlated fashion with a shift towards its pollinator class. Several characters, such as orange/red colour, lack of detectable scent, and a large rostellum are all significantly correlated with bird-pollination. This is in agreement with the classical bird syndrome of Fægri and Van der Pijl (1979) and Johnson (1996b). On the other hand, there is hardly a single character, that is significantly correlated with a shift towards bee-pollination. This imbalance has been pointed out before (Ollerton and Watts 2000) and may be caused by the fact that bird-pollination is, much like wind pollination for instance, a derived state whereas bee-pollination would be a plesiomorphic state.

It is tempting to interpret our results as evidence for the selection of certain floral characters by certain pollinator classes. Our methods were specifically chosen to test this. In the context of natural selection it is important to consider each time that a floral character *could* change, and the number of times that it actually *has* changed under the influence of a certain pollinator class. The STT, unlike the CCT, takes into account the number of branches with characters stasis . In four cases, this resulted in the STT returning a significant value,

whereas the CCT did not. Nonetheless the results should be interpreted with caution. Often the optimizations resulted in floral character changes and pollinator shifts which coincided on a branch. We assumed that the pollinator shift preceeded the floral character change. Therefore our results could simply be an artefact of this assumption. Pagel (1994) described a maximum-likelihood method to explicitly test the order of character evolution (including pollinator shifts) on a phylogeny. Unfortunately, the structure of our dataset did not allow for use of this method, probably because of the small number of observations of some character states and pollinator classes. An alternative explanation for our results would be that the evolution of a certain character state could be considered to be pre-adaptation for a certain pollinator, instead of the result of direct selection (Stebbins 1970; Johnson et al. 1998).

Inferring Pollinator Classes

Given that the accuracy of ancestral character state reconstructions, including pollinator shifts can be compromised by reduced taxon sampling (Salisbury and Kim 2001), it is important to include as many taxa as possible. Typically, if a group of species is large, it is practically impossible to obtain pollinator observations for all species. In the absence of direct pollinator observations, pollinators can be inferred. This is usually based on classical syndrome characters (Fægri and Van der Pijl 1979; e.g. Crisp 1994; Bruneau 1997; Kay et al. 2005). It has been shown, however, that the predictive power of the classical pollination syndromes can be limited (e.g. Herrera 1996; Ollerton 1998). Therefore, statistical methods may offer a better alternative for inferring pollinators. Armbruster (1988) used analysis of covariance and component regression analysis for his numerical continuous data. MDS is a method that is suitable for data including categorical variables such as colour and scent. This was applied by Wilson et al. (2004) to classify species along a ‘bird-bee gradient’ and by Ollerton and Watts (2000) to seek out relationships among the classical pollination syndromes of Fægri and Van der Pijl (1979). We decided to apply this and other methods for inferring pollinators for the unobserved taxa, and to compare the performance among these different methods.

In particular, three observations justified the inference of pollinators. (1) If only the observed species are considered in a MDS analysis, taxa pollinated by the same pollinator class cluster together, almost without overlap. (2) We found several floral characters to evolve in a correlated fashion with certain pollinator classes. (3) The floral characters contained signficantly more structure for grouping taxa pollinated by the same pollinator class compared to random groups. These results together suggest that the floral characters are not randomly

distributed among taxa pollinated by the same pollinator class for the observed species. There is no reason to assume this would be different for the unobserved species.

The method based on a Classification Tree and using the raw data disagreed most strongly with the majority-rule assignment. The reason may be because this method relies on variation among single characters, analogous to identification keys. Therefore, to accurately predict the pollinator class for an unobserved taxon, the value of a character used for prediction *must* fall within the range of values for that particular character among the observed taxa. If this is not the case, for instance due to phylogenetic constraints, the unobserved taxon will possibly be mis-assigned. The Classification Tree that was based on the first three eigenvectors was more congruent with the majority rule assignment. Eigenvectors are compound characters, rather than single characters and they could be regarded as a summary of a whole suite of characters. It is worthwhile to further investigate how one could optimize the Classification Tree method, as it allows for single classes to be split across the tree and, would thereby possibly do more justice to the morphological variation within a pollinator class.

All the other four methods tested had a more or less equally high congruence to the majority rule assignment. The difference between MDS and the other methods is that the other methods force a single pollinator class assignment, whereas with MDS it is more difficult to assign a taxon to a pollinator class based on the ordination (Sakai et al. 1999; Wilson et al. 2004). This ambiguity however, may actually be informative for difficult to place taxa . For example, the bee-pollinated *Satyrium hallackii* ssp. *hallackii*, was placed between the Noctuid+Hawkmoth class and the Bee class in the ordination, not far away from its sister taxon *Satyrium hallackii* ssp. o*cellatum*, which is hawkmoth-pollinated (Johnson 1997a). In this case, both these taxa were observed and therefore assignment was not an issue. There is however no guarantee that such a case could not apply to the unobserved taxa as well. Therefore, MDS is considered the optimal method for inferring pollinators for unobserved taxa for our dataset. Another promising method is to compare the to-class-similarity of unobserved taxa to the among-class-similarity of the observed taxa. Given that the to-class-similarity needs to exceed the among-class-similarity for assignment, it is probably a conservative method.

The success of these methods to assign pollinator classes to unobserved taxa may very well be the consequence of the scale of our pollinator classes. Petanidou and Vokou (1993) and Wilson et al. (2004) showed that it is much more difficult to classify species within a

certain pollinator class. Our results partly demonstrate this with the overlap between the Noctuid and Hawkmoth class in the MDS analysis, even for the observed taxa only.

Our methods for inferring pollinator classes for unobserved taxa suffers from several shortcomings that are mainly due to the fact that unobserved taxa can only be assigned according to the coding of the observed taxa. First of all, unobserved taxa could only be classified according to the pollinator classes that were present among the observed taxa. If pollinator studies are biased towards some pollinators in comparison to others, this may be a problem. This could also apply to rare pollinators which may not be sampled among the observed taxa and therefore cannot be assigned to unobserved taxa either. For our study however, this problem probably does not apply since the pollinator classes among the observed taxa cover almost the entire possible range of pollinator classes (Fægri and Van der Pijl 1979). Furthermore, we assigned each taxon to one pollinator class only. For *Satyrium* this is likely not a problem since all the observed taxa (except for *Satyrium ligulatum*) were pollinated by one pollinator class, and there is no reason to believe this would be different for the unobserved taxa. For other groups this monomorphism may not apply however. Finally, some unrelated pollinator classes may be characterized by similar floral characters (e.g. Beetle and Bee for our study) which could make it difficult to assign an unobserved taxon to either class. This uncertainty could be taken into account by performing the analyses that include inferred pollinator classes for these multiple options. We have taken a different approach by evaluating different methods and assign pollinator classes according to a majority rule criterion.

We recommend that three conditions should be met for inferring pollinators for unobserved taxa. First of all the entire range of floral diversity and phylogenetic diversity should be covered by the observed taxa. Also, the ratio of observed/unobserved taxa should not be too low. For our study this was 0.74. Secondly, a control analysis should be performed to test whether floral characters contain information to group taxa pollinated by the same pollinator class. Thirdly, the proportion of inferred pollinator classes among the unobserved taxa could be contrasted to the proportion of pollinator classes among the observed taxa. In our case, these were rather similar with some exceptions that could be explained by few biases towards choosing taxa for pollinator observations.

Pollinator Shifts

Our results suggest several independent origins of almost all pollinator classes. This seems to be a common finding (e.g. Manning and Linder 1992; McDade 1992; Armbruster

1993; Crisp 1994; Goldblatt and Manning 1996; Bruneau 1997; Hapeman and Inoue 1997; Baum et al. 1998; Johnson et al. 1998; Beardsley et al. 2003;Wilson et al. 2004; Kay et al. 2005; Mant et al. 2005). Nonetheless, the pollinator shifts in *Satyrium* are not completely similar compared to other genera. First of all, the scope of different pollinators is rarely as broad at the species-level as in *Satyrium*. For many other groups, shifts are usually between different pollinators within a certain guild (e.g. Manning and Linder 1992; Armbruster 1993; Bruneau 1997; Mant et al. 2005) or between few pollinator classes, such as bee- and bird-pollination (e.g. McDade 1992; Crisp 1994; Wilson et al.; Kay et al. 2005). Only for the Irid genus *Lapeirousia* (Goldblatt and Manning 1996) and another orchid genus *Disa* (Johnson et al. 1998) was the range of pollinators exploited similarly large. Secondly, the actual number of shifts on the phylogeny is larger for *Satyrium* than it is in most other cases, especially given that the number of shifts in *Satyrium* is underestimated because different pollinators were merged into one class (e.g. Fungus Gnat and Carrion Fly into the Fly class). For *Satyrium* we found a number of 26 shifts out of a maximum possible number of 52 shifts. McDade (1992) found two shifts while the maximum possible number of shifts was 15 for *Aphelandra*. Bruneau (1997) found four shifts out of a maximum possible number of 20 shifts for *Erythrina*. Hapeman and Inoue (1997) found ten shifts out of a maximum possible number of 31 shifts for *Platanthera*. Again, only for *Lapeirousia* (12 out of a maximum possible number of 17 shifts) and *Disa* (17 out of a maximum possible number of 18 shifts) similarly high numbers of shifts were found as for *Satyrium* (Goldblatt and Manning 1996; Johnson et al. 1998). However, sampling of these two genera may have been baised towards sampling as much pollinator diversity as possible. Therefore the number of shifts compared to the maximum number of shifts may actually be slightly lower if all species were included.

The combination of a large number of pollinator shifts among a broad range of pollinators seems to be a unique feature for some southern African lineages. *Satyrium*, as well as *Lapeirousia* and *Disa* have their center of diversity in the Cape Floristic Region (CFR) of South Africa. Even within *Satyrium*, a difference between taxa that occur in the CFR and those outside the CFR was observed. Similarity of floral characters was signficantly smaller among the CFR taxa than among all other taxa (results not shown). Johnson and Steiner (2003) already suggested that a combination of high functional and low species diversity in some pollinator classes may give rise to ecologically specialized plant pollination systems in this region. More southern African genera, that await phylogenetic studies, are also characterized by the presence of multiple pollination systems (Goldblatt and Manning 2006). The special circumstances in southern Africa have implications for nature conservation

because extinction of pollinators may result in a chain-reaction for the many, phylogenetically unrelated, plants relying on them (Johnson and Steiner 2000) and thereby threaten biodiversity.

We found almost no biases in shifts from one pollinator class to another. Only beetle-pollination was significantly more frequently derived from bee-pollination than from any other class. The reason for this may be that these two classes are morphologically very similar and therefore shifts would involve only few morphological changes. In both the Distance and the MDS analysis, the bee- and beetle-pollinated taxa clustered close together. This could be a consequence of the high weighting of the bifid rostellum shape. But also other characters unite taxa pollinated by these two classes, such as the globular viscidium (probably adapted to viscidium placement on the body of the pollinator), short spurs, and nectar production (which separates it from the otherwise similar Fly class). The finding of a tendency to shift from bee- to beetle-pollination is not unique. Armbruster (1993) also found this shift for *Dalechampia* and Steiner (1998) suggested this particular shift for *Ceratandra*, another orchid genus with species in South Africa. The absence of more significant biases may reflect small sample size rather than a random pattern of pollinator shifts. Stebbins (1970) argued that shifts would occur along lines of least resistence. Johnson et al. (1998) found that some shifts were biased towards clades where certain morphological characters had evolved. Upon the collection of more data it may be that in *Satyrium* more biases will be found. Bee-pollination could well be a good candidate for leading to several other pollination classes. Morphologically this pollinator class is intermediate between the classes that are characterized by small or large characters, and bees have a wide distribution range.

Diversification Rates

The hypothesis of a difference in speciation rates among pollinator classes in a phylogenetic framework using molecular branch lengths was tested for the first time in this study. So far, other studies had suggested an increase or decrease in speciation rate with a change in pollination mode (biotic vs. abiotic: Dodd et al. 1999), an increase in speciation rate with a key innovation allowing a lineage to adapt to a wider pollinator spectrum (e.g. Hodges 1997), or the absence of a link between a key innovation and increased speciation rate (von Hagen and Kadereit 2003). Although we found no significant difference in speciation rates among the pollinator classes, the average branch length onto which a certain pollinator class was optimized was different between the pollinator classes. These differences are to some extent expected, especially if the range which pollinators can cover is taken into account.

Certain pollinators probably cover only a small radius resulting in limited pollen-mediated gene flow. As a result, populations can easily become genetically isolated and therefore have a larger potential to speciate (e.g. Barthelmess et al. 2006). It is easy to imagine that birds would be able to cover larger spatial ranges than, for instance, do hawkmoths (Sampson et al. 1995; Johnson et al. 2005). Indeed, we find that the average bird branch is longer than the average hawkmoth branch, suggesting a longer time to speciation for bird-pollinated taxa. The concept may not entirely hold since the putatively least mobile beetles have the longest average branch length. Alternative ways to test for a difference in speciation rates could be to compare the number of species of sister clades (Sanderson and Donoghue 1996). However, our taxon sampling was not sufficient for this method.

Habitat and Pollinators

Taxa pollinated by different pollinator classes seem to be distributed rather equally among habitats, although taxa pollinated by certain pollinator classes are almost confined to specific habitats, for example the fly-pollinated taxa being confined to streamsides, ledges and fynbos, and bird-pollinated taxa being confined to the coastal habitat and grassland. Other habitats are dominated by taxa pollinated by only one pollinator class, such as the noctuid-pollinated taxa in woodland. These findings support the results of Johnson (1997b), who suggested that the pollinator shift between the two subspecies of *Satyrium hallackii* was due to pollinator availability in their respective habitats. However, the absence of a taxon pollinated by a certain pollinator class does not necessarily imply the complete absence of that pollinator class in a habitat. There may also be other reasons for a pollinator shift in a certain habitat, such as interspecific competition for pollinators between taxa within a habitat (Armbruster et al. 1996; Goldblatt and Manning 2006). In fact, the correlation between the number of taxa in a certain habitat and the number of pollinator classes present among these taxa, supports this notion, especially given that many *Satyrium* taxa are sympatric. An association between habitat shifts and pollinator shifts could provide further insight into these two alternative scenarios. If pollinator classes were strongly correlated with certain habitats, we would expect a significant association between habitat shifts and pollinator shifts. If, on the other hand, pollinator classes were present in all habitats, but shifts were to due to interspecific competition, we would expect no pollinator shift in case of habitat shifts and a pollinator shift in the case of no habitat shifts. In our case, we fail to find such associations across the phylogeny. Detailed data on the distribution ranges of both plants and pollinators,

which could have discriminated between some of the processes described above, are unfortunately unavailable.

Topology and Optimization Criterion

Almost all of our results which involved analyses using multiple topologies and optimization criteria are robust to changes in these two variables. This may be because the only differences between the plastid and nuclear topologies concerned a few incongruent taxa (Van der Niet and Linder in review). There was, however, one striking exception to this pattern. For the correlation between genetic distance and similarity of floral characters, the plastid topology returned a significant result, whereas the nuclear topology did not. A plot of the plastid and nuclear genetic distance revealed that a perfect correlation between these two did not exist. After exclusion of the incongruent taxa (Van der Niet and Linder in review), this correlation became linear, demonstrating that the incongruent taxa caused the deviation in correlation. The exclusion of incongruent taxa also resulted in the genetic distance of both the plastid *and* nuclear topology to now return a significant correlation with respect to similarity of floral characters. These results strongly suggest, for the first time to our knowledge, that in case of incongruence the plastid topology tracks morphology better than the nuclear topology does. We interpret this as evidence against the notion that incongruence between plastid and nuclear phylogenies is caused by chloroplast introgression, and that morphology would be naturally tracked by sequences of the nuclear genome (e.g. Okuyama et al. 2005). Our results rather suggest that the incongruence between the plastid and nuclear topologies may either be caused by processes that operate at the molecular level of the nuclear ITS sequences and that are irrelevant to the species tree (Alvarez and Wendel 2003; Van der Niet and Linder in review), or by introgression of neutral nuclear markers that do not track morphological characters that could be under selection (Hodges and Arnold 1994).

Why Pollinator Shifts?

Although our study has demonstrated a high frequency of pollinator shifts and has provided insight into the process of *how* pollinator shifts occur by integrating floral characters, pollinator observations, and habitat data, into a phylogenetic framework, it has not addressed the question *why* pollinator shifts occur. Pollinator shifts can be involved in many biological processes such as avoiding interspecific competition (e.g. Armbruster et al. 1996), evolving reproductive isolation (e.g. Grant 1994), or increasing reproductive output (e.g. Stebbins 1970). Two competing hypotheses exist to explain pollinator shifts for the Cape

Floristic Region in southwestern South Africa, one of the regional centers of diversity of *Satyrium*. Johnson (1996a) suggested that pollinator shifts occur through adaptation to the locally most effective pollinator (Stebbins 1970), while Goldblatt and Manning (1996, 2006) suggested that pollinator shifts are not directly driven by pollinator availability but rather by the evolution of reproductive isolation to select against unfit hybrids through reinforcement (Dobzhansky 1937). Our data do not allow us to distinghuish between these hypotheses. Accurate inference of distribution ranges, measurements of fitness in the field, and more pollinator observations are necessary. However, alternative hypotheses, such as genetic drift (Lande 1976) or pleiotropy, can be rejected based on our data. It would be very difficult to imagine that genetic drift would have resulted in the repeated evolution of certain floral characters in association with the many pollinator shifts, as was the case for *Satyrium*.

ACKNOWLEDGMENTS

We thank Synthesis for providing a grant (BE-TAF-846) to visit the BR herbarium. The National Geographic society, the Swiss Academy of Sciences, and the Georges andAntoine-Claraz Schenkung are thanked for the support provided for fieldwork. Niklaus Zimmermann kindly assisted with the Classification tree analyses. Christo Botes provided data on the pollination of some taxa and assisted in the field. We thank the journal Flora for giving permission to copy Figure 1. Dennis Hansen, Chloé Galley, Merran Matthews and André Simoes provided useful comments on an earlier draft of the manuscript. This work was supported by SNSF grant 31-66594.01, "Adaptive and rapid radiation in the Cape flora".

LITERATURE CITED

Alvarez, I., and J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phyl. Evol. 29: 417-434.

Armbruster, W. S. 1988. Multilevel comparative analysis of the morphology, function, and evolution of *Dalechampia* blossoms. Ecology 69: 1746-1761.

Armbruster, W. S. 1993. Evolution of plant pollination systems: hypotheses and tests with the neotropical vine *Dalechampia*. Evolution 47: 1480-1505.

Armbruster, W. S. 1996. Evolution of floral morphology and function: an integrative approach to adaptation, constraint, and compromise in *Dalechampia* (Euphorbiaceae). Pp. 241-272 in: *Floral biology: studies on floral evolution in animal-pollinated systems,* eds. Lloyd, D. G., and S. C. H. Barrett. New York: Chapman and Hall.

Armbruster, W. S., M. E. Edwards, and E. M. Debevec. 1996. Floral character displacement generates assemblage structure of western Australian triggerplants (*Stylidium*). Ecology 75: 315-329.

Bakker, F. T., L. W. Chatrou, B. Gravendeel, and P. B. Pelser. 2005. *Plant species-level systematics: new perspectives on pattern & process*. Liechtenstein: Gantner Verlag.

Barrett, S. C. H., L. D. Harder, and A. C. Worley. 1996. The comparative biology of pollination and mating in flowering plants. Phil. Trans. R. Soc. Lond. B 351: 1271-1280.

Barthelmess, E. L., C. M. Richards, and D. E. McCauley. 2006. Relative effects of nocturnal vs diurnal pollinators and distance on gene flow in small *Silene alba* populations. New Phytol. 169: 689-698.

Baum, D. A., R. L. Small, J. F. Wendel. 1998. Biogeography and floral evolution of Baobas (*Adansonia*, Bombacaceae) as inferred from multiple data sets. Syst. Biol. 47: 181-207.

Beardsley, P. M., A. Yen, and R. G. Olmstead. 2003. AFLP phylogeny of *Mimulus* section *Erythranthe* and the evolution of hummingbird pollination. Evolution 57: 1397-1410.

Breiman L., J. H. Friedman, R. A. Olshen, C. J. Stone. 1984. *Classification and Regression Trees*. Belmont: Wadsworth International Group.

Bruneau, A. 1997. Evolution and homology of bird pollination syndromes in *Erythrina* (Leguminosae). Am. J. Bot. 84: 54-71.

Crepet, W. L. 1983. The role of insect pollination in the evolution of angiosperms. Pp. 29-50 in: *Pollination biology*, ed. Real, L. New York: Academic Press.

Crepet, W. L. 1984. Advanced (constant) insect pollination mechanisms: pattern of evolution vis-à-vis angiosperm diversity. Ann. Mo. Bot. Gard. 71: 607-630.

Crisp, M. D. 1994. Evolution of bird pollination in some Australian legumes (Fabaceae). Pp. 281-309 in: *Phylogenetics and ecology*, eds. Eggleton, P., and R. I. Vane-Wright. London: Academic Press.

Darwin, C. 1885. *The various contrivances by which orchids are fertilized by insects*. D. Appleton and Co.

Dobzhansky, T. 1937. *Genetics and the origin of species.* New York: Columbia University Press.

Dodd, M. E., J. Silvertown, and M. W. Chase. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. Evolution 53: 732-744.

Dupont, Y. L., D. M. Hansen, J. T. Rasmussen, and J. M. Olesen. 2004. Evolutionary changes in nectar sugar composition associated with switches between bird and insect pollination: the Canarian bird-flower element revisited. Functional Ecology 18: 670-676.

Fægri, K., and L. Van der Pijl. 1979. *The principles of pollination ecology*. Oxford: Pergamon Press.

Fenster, C. B., W. S. Armbruster, P. Wilson, M. R. Dudash, and J. D. Thomson. Pollination syndromes and floral specialization. Annu. Rev. Ecol. Evol. Syst. 35: 375-403.

Fishbein, M., and D. L. Venable. 1996. Diversity and temporal change in the effective pollinators of *Asclepias tuberosa*. Ecology 77: 1061-1073.

Garside, S. 1922. The pollination of *Satyrium bicallosum* Thunb. Ann. Bolus Herb. 2: 149-152.

Goldblatt, P., and J. C. Manning. 1996. Phylogeny and speciation in *Lapeirousia* subgenus *Lapeirousia* (Iridaceae: Ixioideae). Ann. Mo. Bot. Gard. 83: 346-361.

Goldblatt, P., and J. C. Manning. 1998. *Gladiolus in southern Africa.* Vlaeberg: Fernwood Press.

Goldblatt, P., and J. C. Manning. 2006. Radiation of pollination systems in the iridaceae of sub-Saharan Africa. Ann. Bot. 97: 317-344.

Gorelick, R. 2001. Did insect pollination cause increased seed plant diversity? Biol. J. Linn. Soc. 74: 407-427.

Grant, V. 1949. Pollination systems as isolating mechanisms in angiosperms. Evolution 3: 82-97.

Grant, V. 1971. *Plant speciation*. New York: Columbia University Press.

Grant, V. 1994. Modes and origins of mechanical and ethological isolation in angiosperms. Proc. Natl. Acad. Sci. 91: 3-10.

Grant, V., and K. A. Grant. 1965. Flower pollination in the phlox family. New York: Columbia University Press.

Hall, A. V. 1982. A revision of the Southern African species of *Satyrium*. Contributions from the Bolus Herbarium 10: 1-142.

Hapeman, J. R., and K. Inoue. 1997. Plant-pollinator interaction and floral radiation in *Platanthera* (Orchidaceae). Pp. 433-454 in: *Molecular evolution and adaptive radiation*, eds. Givnish, T. J., and K. J. Sytsma. Cambridge: Cambridge University Press.

Harder, L. D., and S. D. Johnson. 1997. Adaptive plasticity of floral display size in animal-pollinated plants. Proc. R. Soc. Lond., B, Biol. Sci. 272: 2651-2657.

Hargreaves, A. L., S. D. Johnson, and E. Nol. 2004. Do floral syndromes predict specialization in plant pollination systems? An experimental test in an "ornitophilous" African Protea. Oecologia 140: 295-301.

Herrera, C. M. 1996. Floral traits and plant adaptation to insect pollinators: a devil's advocate approach. Pp. 65-87 in: *Floral biology: studies on floral evolution in animal-pollinated systems*, eds. Lloyd, D. G., and S. C. H. Barrett. New York: Chapman and Hall.

Hodges, S. A. 1997. Rapid radiation due to a key innovation in columbines (Rannunculaceae: *Aquilegia*). Pp. 391-406 in: *Molecular evolution and adaptive radiation*, eds. Givnish, T. J., and K. J. Sytsma. Cambridge: University Press.

Hodges, S. A., and M. L. Arnold. 1994. Floral and ecological isolation between *Aquilegia formosa* and *Aquilegia pubescens.* Proc. Natl. Acad. Sci. USA 91: 2493-2496.

Hodges, S. A., J. B. Whitall, M. Fulton, and J. Y. Yang. 2002. Genetics of floral traits influencing reproductive isolation between *Aquilegia formosa* and *Aquilegia pubescens*. Am. Nat. 159: S51-S60.

Johnson, S.D. 1996a Pollination, adaptation and speciation models in the Cape flora of South Africa. Taxon 45: 59-66.

Johnson, S. D. 1996b. Bird pollination in south African species of *Satyrium* (Orchidaceae). Plant Syst. Evol. 203: 91-98.

Johnson, S. D. 1997a. Pollination ecotypes of *Satyrium hallackii* (Orchidaceae) in South Africa. Bot. J. Linn. Soc. 123: 225-235.

Johnson, S. D. 1997b. Insect pollination and floral mechanisms in south Africa species of *Satyrium* (Orchidaceae). Plant Syst. Evol. 204: 195-206.

Johnson, S. D., and H. Kurzweil. 1998. Systematics and phylogeny of the *Satyrium erectum* group (Orchidaceae), with descriptions of two new species from the Karoo region of South Africa. Bot. J. Linn. Soc. 127: 179-194.

Johnson, S. D., and K. E. Steiner. 1997. Long-tongued fly pollination and evolution of floral spur length in the *Disa draconis* complex (Orchidaceae). Evolution: 45-53.

Johnson, S. D., and K. E. Steiner. 2000. Generalization versus specialization in plant pollination systems. Trends Ecol. Evol. 15: 140-143.

Johnson, S. D., and K. E. Steiner. 2003. Specialized pollination systems in southern Africa. South African Journal of Science 99: 345-348.

Johnson, S. D., H. P. Linder, and K. E. Steiner. 1998. Phylogeny and radiation of pollination systems in *Disa* (Orchidaceae). Am. J. Bot. 85: 402-411.

Johnson, S. D., P. R. Neal, and L. D. Harder. 2005. Pollen fates and the limits on male reproductive success in an orchid population. Biol. J. Linn. Soc. 86: 175-190.

Kay, K. M., P. A. Reeves, R. G. Olmstead, and D. W. Schemske. 2005. Rapid speciation and the evolution of hummingbird pollination in neotropical *Costus* subgenus *Costus*. Am. J. Bot. 92: 1899-1910.

Kiester, A. R., R. Lande, and D. W. Schemske. 1984. Models of coevolution and speciation in plants and their pollinators. Am. Nat. 124: 220-243.

Kurzweil, H. 1996. Floral morphology and ontogeny in subtribe Satyriinae (Fam. Orchidaceae). Flora 191: 9-28.

Kurzweil, H., and H. P. Linder. 1998. A phylogenetic analysis of the genus *Satyrium* (Orchidaceae). Beiträge zur Biologie der Pflanzen 71: 101-181.

La Croix I. and P. J. Cribb. 1995. Orchidaceae. *Flora Zambesica Vol. 11 part 1*, ed. G. V. Pope. London: Flora Zambesica Managing Committee.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30: 314-334.

Linder, H. P., and H. Kurzweil. 1994. The phylogeny and classification of the Diseae (Orchidoideae: Orchidaceae). Ann. Mo. Bot. Gard. 81: 687-713.

Linder H. P. and H. Kurzweil. 1999. *Orchids of Southern Africa*. Rotterdam: Balkema.

Luckow, M., and A. Bruneau. 1997. Circularity and independence in phylogenetic tests of ecological hypotheses. Cladistics 13: 145-151.

Macior, L. W. 1971. Co-evolution of plants and animals – systematic insights from plant-insect interactions. Taxon 20: 17-28.

Maddison, W. P. 1990. A method for testing the correlated evolution of two binary characters: Are gains and losses concentrated on certain branches of a phylogenetic tree? Evolution 44: 539-557.

Maddison, D. R., and W. P. Maddison. 2005. MacClade 4.07. Sunderland, Massachusetts: Sinauer Associates, Inc.

Manning, J. C., and H. P. Linder. 1992. Pollinators and evolution in *Disperis* (Orchidaceae), or why are there so many species? South African Journal of Science 88: 38-49.

Mant, J., R. Peakall, and P. H. Weston. 2005. Specific pollinator attraction and the diversification of sexually deceptive *Chiloglottis* (Orchidaceae). Plant Syst. Evol. 253: 185-200.

McDade, L. A. 1992. Pollinator relationships, biogeography, and phylogenetics. Bioscience 42: 21-26.

Okuyama, Y., N. Jujii, M. Wakabayashi, A. Kawakita, M. Ito, M. Watanabe, N. Murakami, and M. Kato. Nonuniform concerted evolution and chloroplast capture: heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifracaceae). Mol. Biol. Evol. 22: 285-296.

Ollerton, J. 1996. Reconciling ecological processes with phylogenetic patterns: the apparent paradox of plant-pollinator systems. J. Ecol. 84: 767-769.

Ollerton, J. 1998. Sunbird surprise for syndromes. Nature 394: 726-727.

Ollerton, J., and S. Watts. 2000. Phenotype space and floral typology: Towards an objective assessment of pollination syndromes. Det. Norske Videnskaps – Akademi. I. Matematisk Naturvidenskapelige Klasse, Skrifter, Ny Serie 39: 149-159.

Pagel, M. 1994. Detecting correlated evolution on phylogenies – a general-method for the comparative-analysis of discrete characters. Proc. R. Soc. Lond., B, Biol. Sci. 255: 37-45.

Petanidou, T., and D. Vokou. 1993. Pollination ecology of Labiatae in a phryganic (East Mediterranean) ecosystem. Am. J. Bot. 80: 892-899.

Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817-818.

R Development Core Team. 2006. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Ripley, B. D. 1996. *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.

Rohlf, F. J. 1993. NTSYS-pc Version 1.80. New York: Applied Biostatistics Inc.

Sakai, S., M. Kato, and T. Inoue. 1999. Three pollination guilds and variation in floral characteristics of Bornean gingers (Zingiberaceae and Costaceae). Am. J. Bot. 86: 646-658.

Salisbury, B. A., and J. H. Kim. 2001. Ancestral state estimation and taxon sampling density. Syst. Biol. 50: 557-654.

Sampson, J. F., S. D. Hopper, and S. H. James. 1995. The mating system and genetic diversity of the Australian arid zone mallee, *Eucalyptus rameliana*. Aust. J. Bot. 43: 461-474.

Sanderson, M. J., and M. J. Donoghue. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. Trends Ecol. Evol. 11: 15-20.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: a Penalized Likelihood approach. Mol. Biol. Evol. 19: 101-109.

Sanderson, M. J. 2003. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19: 301-302.

Schemske, D. W., and H. D. Bradshaw. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). Proc. Natl. Acad. Sci. U.S.A. 96: 11910-11915.

Sillén-Tullberg, B., 1993. The effect of biased inclusion of taxa on the correlation between discrete characters in phylogenetic trees. Evolution 47: 1182-1191.

Stebbins jr., G. L. 1970. Adaptive radiation of reproductive characteristics in angiosperms I: pollination mechanisms. Annu. Rev. Ecol. Syst. 1: 307-326.

Steiner, K. E. 1998. The evolution of beetle pollination in a South African orchid. Am. J. Bot. 85: 1180-1193.

Summerhayes V. S. 1968. Orchidaceae (part 1). *Flora of tropical East Africa*, eds. E. Milne-Redhead and R. M. Polhill. London: Crown Agents for Oversea Governments and Adminstrations.

Swofford, D.L. 2000. PAUP*: Phylogenetic Analysis Using Parsimony, version 4.0b. Sinauer Associates, Sunderland, Massachusetts.

Van der Niet, T., H. P. Linder, B. Bytebier, and D. U. Bellstedt. 2005. Molecular markers reject monophyly of the subgenera of *Satyrium* (Orchidaceae). Syst. Bot. 30:263-274.

Van der Pijl., L., and C. H. Dodson. 1966). *Orchid flowers; their pollination and evolution.* Coral Gables, Florida: The Fairchild Tropical Garden and the University of Miami Press.

Von Hagen, K. B., and J. W. Kadereit. 2003. The diversification of *Halenia* (Gentianaceae): ecological opportunity versus key innovation. Evolution 57: 2507-2518.

Vogel, S. 1954. Blütenbiologische typen als elemente der sippengliederung. Bot. Stud. 1: 1-337.

Vogel, S. 1959. Organographie der Blüten kapländischer Ophrydeen. Adak. Wiss. Mainz, Abh. Math.-Naturwiss. Kl. 6-7: 1-268.

Waser, M. W. 1998. Pollination, angiosperm speciation, and the nature of species boundaries. Oikos 81: 198-201.

Waser, N. M., L. Chittka, M. V. Price, N. Williams, J. Ollerton. 1996. Generalization in pollination systems, and why it matters. Ecology 77: 279-296.

Weller , S. G., and A. K. Sakai. 1999. Using phylogenetic approaches for the analysis of plant breeding system evolution. Annu. Rev. Ecol. Syst. 30: 167-199.

Wilson, P., and J. D. Thomson. 1996. How do flowers diverge? Pp. 88-111 in: *Floral biology: studies on floral evolution in animal-pollinated systems*, eds. Lloyd, D. G., and S. C. H. Barrett. New York: Chapman and Hall.

Wilson, P., M. C. Castellanos, J. N. Hogue, J. D. Thomson, and W. S. Armbruster. 2004. A multivariate search for pollination syndromes among penstemons. Oikos 104: 345-361.

Wilson, P., M. C. Castallanos, A. D. Wolfe, and J. D. Thomson. 2006. Shifts between bee and bird pollination in Penstemons. Pp. 47-68 in: *Plant-Pollinator interactions: From specialisation to generalisation*, eds. Waser, N. M., and J. Ollerton. Chicago: The University of Chicago Press.

Macro-Evolutionary Data suggest a Role for Reinforcement in Pollination System Shifts

TIMOTHEÜS VAN DER NIET[1], STEVEN D. JOHNSON[2], AND H. PETER LINDER[1]

[1]*Institute of Systematic Botany, University of Zurich, Zollikerstrasse 107, 8008, Zurich, Switzerland*

[2]*School of Biological and Conservation Sciences, University of KwaZulu-Natal, P. Bag X01 Scottsville, Pietermaritzburg 3209, South Africa*

Submitted to Evolution

Abstract.—Reproductive isolation can evolve either as a by-product of divergent selection, or through reinforcement. We used the Cape flora of South Africa, known for its high level of pollination specialisation, as a model system to test the potential role of pollinator-mediated selection in the speciation process. Comparative analysis of 41 sister-species pairs (representing Geraniaceae, Iridaceae, and Orchidaceae) for which complete pollinator, edaphic and distribution data are available showed that for sister species with overlapping distribution ranges, pollination system shifts are significantly associated with edaphic shifts. In contrast, there is no significant association between pollination system shifts and edaphic shifts for allopatric sister species. These results are interpreted as evidence for reinforcement.

Keywords: speciation, pollination, plants, by-product, reinforcement, Cape flora, distribution ranges

INTRODUCTION

One of the greatest challenges in the study of ecological speciation is to distinguish the ways in which divergent selection has led to reproductive isolation (Schluter 2001). One hypothesis predicts that reproductive isolation evolves as a by-product of adaptation in response to divergent selection operating in different environments (Mayr 1942). This idea is supported by extensive empirical data (Rice and Hostert 1993; McKinnon et al. 2004; Rieseberg et al. 2004; Orr 2005). A competing hypothesis predicts that reproductive isolation is directly selected for. This process of reinforcement operates when incipient species come into secondary contact, and is the result of selection against inferior hybrid offspring (Dobzhansky 1937). The concept of reinforcement was initially dismissed in earlier studies (Butlin 1989; Rice and Hostert 1993), however it now appears that, at least in theory, reinforcement is plausible (Noor 1999; Servedio and Noor 2003). Several empirical studies suggest that reinforcement may operate in nature (Coyne and Orr 1989; Hoskin et al. 2005; Lukhtanov et al. 2005), although distinguishing between reinforcement and alternative processes, especially those where reproductive character displacement can be attributed to causes other than the prevention of heterospecific matings is difficult (Noor 1999).

The vast majority of studies addressing the origin of reproductive isolation use animals as their model system (Rice and Hostert 1993; Noor 1999; Schluter 2001; Orr 2005). Grant (1994) claimed however, that the same two mechanisms for evolution of reproductive isolation also operate in plants. In several plant genera, reproductive isolation is thought to have evolved as a by-product of adaptation to different pollinators (Grant 1993; Schemske and Bradshaw 1999) and different edaphic conditions (Macnair and Christie 1983). Evidence for reinforcement in plants comes from the observation of greater flower colour differences between sympatric populations, as compared to allopatric populations of different species (Levin and Kerster 1967; Levin 1978), plus the evolution of selfing in response to proximity of a congener, resulting in reproductive isolation (Antonovics 1967; Fishman and Wyatt 1999).

Although useful, most of these case studies lack a phylogenetic framework and are therefore unable to identify sister species. By restricting the investigation to sister species, the chance that observed differences and their underlying processes are involved in speciation, is maximized. This is particularly important for studying reinforcement, as reproductive character displacement is not restricted to reinforcement alone, but could also occur in response to competition, for instance, for pollinators (Armbruster et al. 1994). Moyle et al. (2004) used a phylogenetic framework to investigate patterns of reproductive isolation in

three angiosperm genera. Unfortunately, their study only takes post-pollination stages into account and therefore has limited power in studying the speciation process in general, and reinforcement in particular.

The Cape Floristic Region (CFR) in southern Africa provides an excellent model system to address hypotheses for the evolution of reproductive isolation. With its high species richness and levels of endemism it is thought to have been an arena for speciation (Goldblatt 1978; Goldblatt and Manning 2002; Linder 2003, 2005). Four observations suggest that shifts in pollination system may have played a role in the speciation process: (i) many large and florally diverse genera exist in the CFR (Johnson 1996b; Linder 2003), (ii) the region is characterized by a remarkably high number of specialized pollination systems (Johnson and Steiner 2003), (iii) several pollination systems often occur within groups of closely related species (e.g. Johnson et al. 1998; Manning and Goldblatt 2005), and (iv) there is evidence for strong selection by pollinators (Johnson and Steiner 1997) resulting in convergent adaptations (e.g. Goldblatt and Manning 2000).

Johnson (1996b) suggested that floral shifts and associated adaptive shifts in pollination system were most probably a response to selection for increased efficiency of the female function of flowers in the pollen-limited CFR. Given that shifts in pollination system may result in reproductive isolation, this hypothesis implies that reproductive isolation evolves as a by-product of optimizing female floral function. In contrast, Goldblatt and Manning (1996, 1998) suggested that shifts in pollination system occur upon secondary contact to prevent gene flow between incipient species that first diverged on different, but often adjacent soils. The evolution of reproductive isolation by a shift in pollination system is then directly selected for, and therefore this hypothesis is consistent with reinforcement.

Here we test these competing hypotheses for the origin of reproductive isolation in the Cape flora by evaluating shifts in pollination system between sister plant species. More specifically, we test for an association between pollination system shifts and edaphic shifts for sister species with overlapping distribution ranges and allopatric sister species respectively. Furthermore we explicitly test for an association between a joint shift in both pollination system and edaphic conditions and overlapping distribution ranges. If a shift in pollination system evolves as a by-product of direct adaptation to locally effective pollinators (Stebbins 1970), we predict that there will be no association between pollination system shifts and edaphic shifts, regardless of distribution ranges. If pollination system shifts occur through reinforcement, we predict an association between pollination system shifts and edaphic conditions for sister species with overlapping distribution ranges, but *not* for allopatric sister

species. Furthermore we predict that a joint shift in both pollination system and edaphic conditions is associated with overlapping distribution ranges.

METHODS

For the following six genera a sufficiently sampled and resolved phylogeny to confidently identify sister species, plus data on pollination systems, edaphic attributes, and distribution ranges were available: (*Pelargonium* in the Geraniaceae, *Gladiolus* and *Lapeirousia* in the Iridaceae, and *Ceratandra*, *Disa* and *Satyrium* in the Orchidaceae). Forty-one sister-species pairs were selected from the published phylogenies and their attributes were taken from the literature (Table 1). Distribution ranges were scored as 'overlapping' if sister species occurred within the same quarter degree grid for at least part of their distribution ranges.

TABLE 1. Selected genera for this study with their source in the literature, including number of sister-species comparisons (N) and some of the key attributes.

genus	Pollination system shift	edaphic shift	Pollination system + edaphic shift	allopatry/overlapping distribution range
Pelargonium (N=6)[a]	3	4	3	2/4
Gladiolus (N=16)[b]	6	7	3	9/7
Lapeirousia (N=6)[c]	3	4	2	3/3
Ceratandra (N=2)[d]	1	1	1	1/1
Disa (N=6)[e]	5	3	3	1/5
Satyrium (N=5)[f]	4	2	2	2/3

[a] Phylogeny taken from Bakker et al. (2004), pollinator data were inferred from floral syndromes (Struck 1997; M. Struck, pers. com., 2005), soil and distribution data were taken from Van der Walt (1985), Van der Walt and Boucher (1986), Van der Walt and van Zyl (1988), Albers et al. (2000), and Marais (2005).
[b] Phylogeny, pollinator, soil, and distribution data taken from Goldblatt and Manning (1998). The phylogeny, based on morphological characters, is not reconstructed using a rigorous cladistic method.
[c] Phylogeny, pollinator, soil, and distribution data taken from Goldblatt and Manning (1996).
[d] Phylogeny, pollinator, soil, and distribution data taken from Linder and Kurzweil (1999).
[e] Phylogeny and pollinator data taken from Johnson et al. (1998), soil and distribution data taken from Linder and Kurzweil (1999).
[f] Phylogeny taken from Van der Niet (unpublished) and Johnson and Kurzweil (1998), pollinator data taken from Johnson (1996a, 1997, unpublished data), soil and distribution data taken from Linder and Kurzweil (1999).

The frequency of shifts in pollination system and edaphic conditions was determined for all available sister-species pairs. A shift in pollination system between members of a pair was scored when the species differed in either their pollinators or in their breeding system (e.g. animal pollinated vs. autogamy). A shift in edaphic conditions was interpreted as either a difference in soil type (e.g. sandstone versus clay) or soil moisture (e.g. dry versus swampy habitat).

In order to determine whether shifts in pollination system were more frequent among sister species that had undergone an edaphic shift in comparison to those that had not, we used a one-tailed Fisher Exact test. This test was performed for sister species with overlapping distribution ranges and allopatric sister species, respectively. We used the G test for heterogeneity of counts applying Williams correction (Sokal and Rohlf 1981) to explore associations between a joint shift in pollination system and edaphic conditions, and geographical distribution range (allopatric versus overlapping).

RESULTS

Of the 41 sister-species pairs, a shift in pollination system was found for 22 of them. Twenty-three sister-species pairs had overlapping distribution ranges. Of the 12 sister species with overlapping distribution ranges that showed a shift in edaphic conditions, 11 (92%) pairs also involved a shift in pollination system. Only five (45%) of the sister species with overlapping distribution ranges showed a shift in pollination system in the absence of a shift in edaphic conditions (Fig. 1). Therefore, a shift in pollination system is significantly ($\alpha<0.05$) associated with a shift in edaphic conditions ($p=0.024$, one-tailed Fisher Exact test). This association was not significant for the 18 sister-species pairs with allopatric distribution ranges (Fig. 1) ($p=0.69$, one-tailed Fisher Exact test). Species that have undergone a joint shift in pollination system and edaphic conditions were significantly ($\alpha<0.05$) associated with overlapping distribution ranges compared to species that either shifted in one or none of these ecological variables (Fig. 1) ($\chi^2=4.39$, d.f.=1, $p=0.036$).

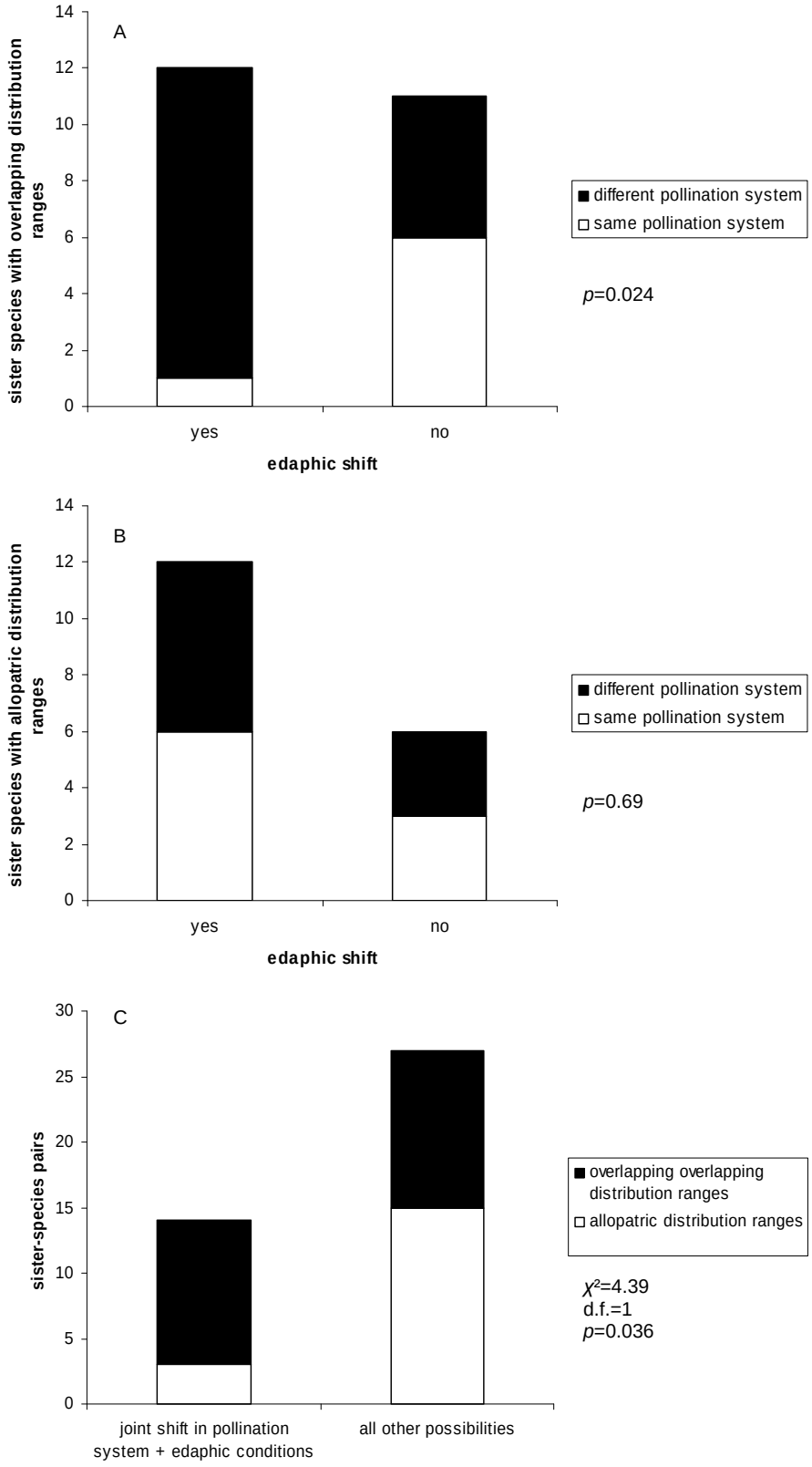


FIG. 1. a. The association of pollination system shifts and edaphic conditions for sister species with overlapping distribution ranges. b. The association of pollination system shifts and edaphic conditions for allopatric sister species. c. The association between a joint shift in both pollination system and edaphic conditions and distribution range for all sister species.

DISCUSSION

Our macro-evolutionary data suggest that shifts in pollination system are associated with shifts in physical habitat (edaphic conditions) for sister species with overlapping distribution ranges. At the same time, this association appears to be absent for sister species with allopatric distribution ranges. Finally, there seems to be an association between overlapping distribution ranges and a joint shift in both pollination system and edaphic conditions. These results are consistent with the idea of reinforcement whereby incipient species initially diverge on different soils and, upon secondary contact, adapt to different pollinators as the result of selection against unfit hybrids (Goldblatt and Manning 1996, 1998). The alternative hypothesis, local adaptation to effective pollinators (Grant and Grant 1965; Johnson 1996b), does not predict the observed association between pollination system shifts, edaphic shifts and overlapping distribution ranges. There is ample evidence to suggest that edaphic factors are of great importance for plant speciation in general (Rajakaruna 2004) and in the Cape flora in particular (Kurzweil et al. 1991). Indeed, one of the strongest cases for reinforcement in plants comes from studies on species that are adapted to soils that are contaminated with heavy metals (Antonovics 1967).

Some intrinsic features of the CFR may facilitate the presence of reinforcement. The region is characterized by a mosaic of different habitats that are sharply distinct but not well separated in space. This applies particularly to different soil types (Cowling 1992). That these conditions affect plant distribution is suggested by the high turnover of species along habitat gradients in the CFR (Linder 1985). At the same time, the level of specialized pollination systems is remarkably high in the CFR (Johnson and Steiner 2003). This high level could relax the constraint on adaptation to different pollinators, which is usually imposed on species with a generalist pollination system (Waser 1998).

Our study is limited to biotically pollinated sister-species pairs whereas the CFR is also rich in wind-pollinated lineages (Linder 2003). However, there is no reason to believe that reinforcement can only occur through shifts in pollination system. Phenological shifts could equally well result in reinforcement (Antonovics 1967; Silvertown et al. 2005). Indeed, in the CFR, several sister species of Restionaceae flower in different seasons (Linder 2003).

Several alternative processes which could also result in a joint shift in pollination system and edaphic conditions were rejected by our data. Firstly, in contrast to reinforcement where adaptation to different edaphic conditions should precede a shift in pollination system, cladogenesis could have occurred through adaptation to a new pollinator with subsequent anagenetic adaptation to different edaphic conditions. However, we would then expect to find a similar association for allopatric sister species, which we did not find. In addition we would

not expect a joint shift in both pollination system and edaphic conditions to be associated with overlapping distribution ranges. Secondly, pollinator distribution may be partly linked to soil, given that edaphic conditions are often an important determinant of vegetation. Many pollinator species rely on vegetation or edaphic conditions for nesting sites, feeding and protection (e.g. Eyre and Luff 2005). Shifts in pollinators and soils may thus occur as parallel processes. However, again, we would then expect this association to be equally likely for allopatric sister species, and we would not expect the association between distribution ranges and a joint shift in pollination system and edaphic conditions. In addition, we found sister species that shifted in pollination system but not in edaphic conditions. Finally, reproductive character displacement between taxa with overlapping distribution ranges, possibly as a result of competition for the same pollinator, could lead to shifts in pollination system after speciation in response to selection for different edaphic conditions. This process can leave a similar signature as would reinforcement. Therefore it is difficult to discriminate between these two processes. However, we still believe that reinforcement is a more likely explanation of our results for two reasons. Firstly, we have used sister species of genera that are known to hybridize in nature (Hall 1982; van der Walt 1990; Ellis and Johnson 1999), a condition which is necessary for reinforcement. Secondly, and more importantly, while reproductive character displacement is thought to be most likely among taxa that occur in full sympatry with no divergence in physical habitat (Armbruster et al. 1994), reinforcement requires some kind of divergence between incipient species, which results in maladapted hybrids (Dobzhansky 1937). We found that the majority of sister species with overlapping distribution ranges that shifted in pollination system were found on different soil types as well.

A role for reinforcement in plant speciation has been long suggested (e.g. Levin 1978). In this study we test for the first time its role on a large scale. Our results are consistent with the hypothesis that reinforcement occurs through the achievement of reproductive isolation by a shift in pollination system. However, further studies are needed that focus on identifying how reinforcement could occur, including the establishment of hybrid fitness and what influences it, the evaluation of features that are important in pollinator attraction for ranges of overlap and allopatry between sister species, and comparison with abiotically pollinated plant species.

ACKNOWLEDGMENTS

We thank M. Struck for providing us data on pollination syndromes of *Pelargonium*, Merran Matthews, Bruce Anderson, Allen Ellis, and Florian Schiestl for his helpful comments on an earlier draft of the manuscript.

LITERATURE CITED

Albers, F., J .J. A. van der Walt, E. M. Marais, M. Gibby, D. E. Marschewski, A. M. van der Merwe, F. T. Bakker, and A. Culham. 2000. A biosystematic study of *Pelargonium* section *Ligularia*: 4. The section *Ligularia* sensu stricto. S. Afr. J. B. 66: 31-43.

Antonovics, J. 1967. Evolution in closely adjacent plant populations. V. Evolution of self-fertility. Heredity 23: 219-238.

Armbruster, W. S., M. E. Edwards, and E. M. Debevec. 1994. Floral character displacement generates assemblage structure of Western Australian triggerplants (*Stylidium*). Ecology 75: 315-329.

Bakker, F. T., A. Culham, P. Hettiarachi, T. Touloumenidou, and M. Gibby. 2004. Phylogeny of *Pelargonium* (Geraniaceae) based on DNA sequences from three genomes. Taxon 53: 17-28.

Butlin, R. 1989. Reinforcement of premating isolation. Pp. 158-179 in: *Speciation and its consequences* eds. Otte D., and J.A. Endler, eds.. Sunderland: Sinauer Associates.

Coyne, J. A. and H. A. Orr. 1989. Patterns of speciation in *Drosophila*. Evolution 43: 362-381.

Cowling, R. M. 1992. *The ecology of fynbos*. Cape Town: Oxford University Press.

Dobzhansky, T. 1937. *Genetics and the origin of species.* New York: Columbia University Press.

Ellis, A. G. and S. D. Johnson. 1999. Do pollinators determine hybridization patterns in sympatric *Satyrium* (Orchidaceae) species? Plant Syst. Evol. 219: 137-150.

Eyre, M. D. and M. L. Luff. 2005. The distribution of Epigeal Beetle (Coleoptera assemblages on the north-east England Coast. Journal of Coastal Research 21: 982-990.

Fishman, L. and R. Wyatt. 1999. Pollinator-mediated competition, reproductive character displacement, and the evolution of selfing in *Arenaria uniflora* (Caryophyllaceae). Am. J. Bot. 53: 1723-1733.

Goldblatt, P. 1978. An analysis of the flora of Southern Africa: its characteristics, relationships, and origins. Ann. Mo. Bot. Gard. 65: 363-436.

Goldblatt, P. and J. C. Manning. 1996. Phylogeny and speciation in *Lapeirousia* subgenus *Lapeirousia* (Iridaceae: Ixioideae). Ann. Mo. Bot. Gard. 83: 346-361.

Goldblatt, P. and J. C. Manning. 1998. *Gladiolus in southern Africa*. Vlaeberg: Fernwood Press.

Goldblatt, P. and J. C. Manning. 2000. The long-proboscid fly pollination system in southern Africa. Ann. Mo. Bot. Gard. 87: 146-170.

Goldblatt, P. and J. C. Manning. 2002. Plant diversity of the Cape Region of Southern Africa. Ann. Mo. Bot. Gard 89: 281-302.

Grant, V. 1993. Origin of floral isolation between ornitophilous and sphingophilous plant species. Proc. Natl. Acad. Sci. U.S.A. 90: 7729-7733.

Grant, V. 1994. Modes and origins of mechanical and ethological isolation in angiosperms. Proc. Natl. Acad. Sci. U.S.A. 91: 3-10.

Grant, V. and K.A. Grant. 1965. *Flower pollination in the Phlox family.* New York: Columbia University Press.

Hall, A.V. 1982. A revision of the Southern African species of Satyrium. Contributions from the Bolus Herbarium 10: 1-142.

Hoskin, J., M. Higgie, K. R. McDonald, and C. Moritz. 2005. Reinforcement drives rapid allopatric speciation. Nature 437: 1353-1356.

Johnson, S. D. 1996a. Bird pollination in South African species of *Satyrium* (Orchidaceae). Plant Syst. Evol. 203: 91-98.

Johnson, S. D. 1996b. Pollination, adaptation and speciation models in the Cape flora of South Africa. Taxon 45: 59-66.

Johnson, S. D. 1997. Insect pollination and floral mechanisms in South African species of *Satyrium* (Orchidaceae). Plant Syst. Evol. 204: 195-206.

Johnson, S. D. and W. J. Bond. 1992. Habitat dependent pollination success in a cape orchid. Oecologia 91: 455-456.

Johnson, S. D. and H. Kurzweil. 1998. Systematics and phylogeny of the *Satyrium erectum* group (Orchidaceae), with description of two new species from the Karoo region of South Africa. Bot. J. Linn. Soc. 127: 179-194.

Johnson, S. D. and K. E. Steiner. 1997. Long-tongued fly pollination and evolution of floral spur length in the *Disa draconis* complex (Orchidaceae). Evolution 51: 45-53.

Johnson, S. D. and K. E. Steiner. 2003. Specialized pollination systems in southern Africa. S. Afr. J. Sci. 99: 345-348.

Johnson S. D., H. P. Linder, and K. E. Steiner. 1998. Phylogeny and radiation of pollination systems in *Disa* (Orchidaceae). Am. J. Bot. 85: 402-411.

Kurzweil, H., H. P. Linder, and P. Chesselet. 1991. The phylogeny and evolution of the *Pterygodium* - *Corycium* complex (Coryciinae, Orchidaceae). Plant Syst.Evol. 175: 161-223.

Levin, D. A. 1978. The origin of isolating mechanisms in flowering plants. Evol. Biol. 11: 185-317.

Levin, D. A. and H. W. Kerster. 1967. Natural selection for reproductive isolation in *Phlox*. Evolution 21: 679-687.

Linder, H. P. 1985. Gene flow, speciation and species diversity patterns in a species rich area: the Cape flora. Pp. 53-57 in: *Species and speciation*, ed. Vrba, E. D. Pretoria: Transvaal Museum Monographs.

Linder, H. P. 2003. The radiation of the cape flora, southern Africa. Biol. Rev. 78: 597-638.

Linder H. P. 2005. The evolution of diversity: the Cape flora. Trends Plant Sci. 10: 536-541.

Linder, H. P. and H. Kurzweil. 1999. *Orchids* of southern Africa. Balkema, Rotterdam.

Lukhtanov, V. A., N. P. Kandul, J. B. Plotkin, A. V. Dantchenko, D. Haig, D. and N. E. Pierce. 2005. Reinforcement of pre-zygotic isolation and karyotype evolution in *Agrodiaetus* butterflies. Nature 436: 385-389.

Macnair, M. R. and P. Christie. 1983. Reproductive isolation as a pleiotropic effect of copper tolerance in *Mimulus guttatus*? Heredity 50: 295-302.

Manning, J. C. and P. Goldblatt. 2005. Radiation of pollination systems in the Cape genus Tritoniopsis (Iridaceae: Crocoideae) and the development of bimodal pollination strategies. Int. J. Plant Sci. 166: 459-474.

Marais, E. M. 2005. Differences between *Pelargonium moniliforme* (Geraniaceae) and the closely related *P. vinaceum*. S. Afr. J. Bot. 71: 221-227.

Mayr, E. 1942 Systematics and the origin of species.Columbia University Press, New York.

McKinnon, J. S., S. Mori, B. K. Blackman, L. David, D. M. Kingsley, L. Jamieson, J. Chou, and D. Schluter. 2004. Evidence for ecology's role in speciation. Nature 429: 294-298.

Moyle, L. C., M. S. Olson, and P. Tiffin. 2004. Patterns of reproductive isolation in three angiosperm genera. Evolution 58: 1195-1208.

Noor, M. A. F. 1999. Reinforcement and other consequences of sympatry. Heredity 83: 503-508.

Orr, H. A. 2005. The genetic basis of reproductive isolation: Insights from *Drosophila*. Proc. Natl. Acad. Sci. U.S.A. 102: 6522-6526.

Rajakaruna, N. 2004. The edaphic factor in the origin of plant species. Int. Geol. Rev. 46: 471-478.

Rice, W. R. and E. E. Hostert. 1993. Laboratory experiments on speciation: what have we learned in 40 years? Evolution 47: 1637-1653.

Rieseberg, L. H., S. A. Church, and C. L. Morjan. 2004. Integration of populations and differentiation of species. New Phytol. 161: 59-69.

Schemske, D. W. and H. D. Bradshaw Jr. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). Proc. Natl. Acad. Sci. U.S.A. 96: 11910-11915.

Schluter, D. 2001. Ecology and the origin of species. Trends Ecol. Evol. 16: 372-380.

Servedio, M. R. and M. A. F. Noor. 2003. The role of reinforcement in speciation: Theory and data. Annu. Rev. Ecol. Syst. 34: 339-364.

Silvertown, J., C. Servaes, P. Biss, and D. Macleod. 2005. Reinforcement of reproductive isolation between adjacent populations in the park grass experiment. Heredity 95, 198-205.

Sokal, R. R. and F. J. Rohlf. 1981. *Biometry*. New York: W.H. Freeman and Company.

Stebbins Jr., G. L. 1970. Adaptive radiation of reproductive characteristics in angiosperms, I: pollination mechanisms. Annu. Rev. Ecol. Syst. 1: 307-326.

Struck, M. 1997. Floral divergence and convergence in the genus *Pelargonium* (Geraniaceae) in southern Africa: ecological and evolutionary considerations. Plant Syst. Evol. 208: 71-97.

Van der Walt, J. J. A. 1985. A taxonomic revision of the type section of *Pelargonium* L'Hérit. (Geraniaceae). Bothalia 15: 345-386.

Van der Walt, J. J. A. and D. A. Boucher. 1986 A taxonomic revision of the section *Myrrhidium* of *Pelargonium* (Geraniaceae) in southern Africa. S. Afr. J. Bot. 52: 438-462.

Van der Walt, J. J. A. and L. van Zyl. (nee Hugo) 1988. A taxonomic revision of *Pelargonium* section *Campylia* (Geraniaceae). S. Afr. J. Bot. 54: 145-171.

Van der Walt, J. J. A., F. Albers, and M. Gibby. 1990. Delimitation of *Pelargonium* sect. *Glaucophyllum* (Geraniaceae). Plant Syst. Evol. 171: 15-26.

Waser, N. M. 1998. Pollination, angiosperm speciation, and the nature of species boundaries. Oikos 82: 198-201.

SUMMARY

In the first three chapters of this thesis I present the results of a macro-evolutionary study where I address three main issues on congruence between taxonomy and phylogenetic relationships, reconstruction of a species tree, and what factors may have driven diversification.

Infrageneric Classification

In chapter 1 I tested the monophyly of the subgenera of *Satyrium*. Although the monophyly of *Satyrium* has been established using both morphological as well as molecular data, the infrageneric classification has been highly problematic due to the large morphological diversity. Therefore several different classifications are available. The only infrageneric classification based on a phylogenetic analysis of morphological characters resulted in recognition of the three subgenera *Brachysaccium*, *Bifidum* and *Satyrium*. DNA sequence data from the nuclear (ITS1, 5.8S and ITS2) and plastid (*trnL* intron and *trnL-F* intergenic spacer and part of the *matK* gene and *trnK* intron) genome were used to test the monophyly of these subgenera, and the status of the aberrant species *S. rhynchanthum* and *S. pumilum*. A combined phylogenetic analysis was performed on a dataset from which two taxa that were incongruent between the plastid and nuclear topology were pruned. *S. rhynchanthum* and *S. pumilum* were both found to be nested within *Satyrium*. Parametric bootstrap, as well as Bayesian posterior probability, rejected monophyly of all three subgenera and alternative groupings are suggested instead. Several morphological characters were optimized onto the phylogeny which distinghuish the five main clades that were suggested by the molecular data. Not a single character could be found that is either consistent among members of a clade or useful for identification.

The Species Tree

In chapter 2 I describe a detailed protocol to deal with phylogenetic incongruence in the quest for the species tree. This involved three steps: identifying incongruence, assessing the cause of incongruence, and reconstructing the species tree. Separate phylogenetic analysis of a plastid and nuclear dataset for 63 species of *Satyrium* revealed many cases of incongruence. The Incongruence Length Difference test showed that many of these were, in fact, non-significant. For the remaining significant cases, results from taxon jack-knife experiments and parametric bootstrap suggested that non-biological artefacts such as sparse taxon sampling and long-branch attraction could be excluded as causes for the observed

incongruence. In order to evaluate biological causes, such as orthology/paralogy conflation, lineage sorting, and hybridization, the number of events was counted that need to be invoked *a-posteriori* to explain the observed pattern. In most cases where incongruence was significant, this resulted in an almost equal number of events for each of these different causes. Only for the three species from south east Asia, that form a monophyletic clade, hybridization was favoured over the alternative causes. This conclusion is based on the large number of events that needs to be invoked, in order for either orthology/paralogy conflation or lineage sorting to have been the cause of the incongruence. Morphological evidence further supports a hybrid origin of this clade. The final species tree presented is the product of the combined analysis of both plastid and ribosomal nuclear DNA sequences for all congruent species and *a-posteriori* grafting of the incongruent clades or accessions onto the tree. This tree provides the best phylogenetic hypothesis to date, and serves as a template for subsequent evolutionary studies.

Pollinator Shifts in Satyrium

In chapter 3 I investigated the interaction between floral characters and pollinator shifts in a phylogenetic framework. For 25 taxa of *Satyrium* pollinator observations are available. These observations show that almost all taxa are solely pollinated by a single pollinator class, including bee-, bird-, butterfly-, fungus gnat-, beetle-, carrion fly-, noctuid-, and hawkmoth-pollination. Floral characters which are putatively involved in pollination contained information to group taxa that are pollinated by the same pollinator class together. The floral characters also showed significant phylogenetic structure. However, although the amount of similarity of floral characters was negatively correlated with genetic distance, it was significantly greater among taxa that are pollinated by the same pollinator class than among members of a clade. As expected, genetic distance was significantly greater among members of a clade than among taxa that are pollinated by the same pollinator class. We found certain floral characters that evolved in a correlated fashion with shifts to a certain pollinator class, most notably with bird-pollination. To increase sample size we inferred pollinator classes for another 35 taxa by using and evaluating different methods such as Non-metric Multidimensional Scaling, Distance analysis, Classification Trees, and various measurements and comparisons of similarity of floral characters. We found that the final assignment of a pollinator class to the unobserved taxa was most often according to the assignment of the Non-metric Multidimensional Scaling, and least often according to the Classification Tree. Optimization of pollinator classes onto a species-level phylogeny

revealed that all pollinator classes have multiple origins. We observed that the number of pollinator shifts on the phylogeny is 50% of the maximum possible number of shifts given the data. Apart from a bias from bee- to beetle-pollination, there was no pattern of directionality among the shifts. We found no significant difference in speciation rates among the different pollinator classes. Taxa pollinated by a certain pollinator class seem to be distributed across habitats with little biases. We found a positive relationship between the number of taxa ocurring in a certain habitat and the number of pollinator classes among these taxa. We did not find any assocation between habitat- and pollinator shifts. We interpret our results as evidence that the evolution of floral characters is dependant on the phylogeny, but more so on the pollinators. The number of pollinator shifts is high compared to other studies. This may be related to the specific pollination conditions in southern Africa.

Shifts in Pollination System in the Cape flora

In chapter 4 I focused on the evolution of reproductive isolation in the Cape flora. Reproductive isolation can evolve either as a by-product of divergent selection, or through reinforcement. We used the Cape flora of South Africa, known for its high level of pollination specialisation, as a model system to test the potential role of pollinator-mediated selection in the speciation process. Comparative analysis of 41 sister-species pairs (representing Geraniaceae, Iridaceae, and Orchidaceae) for which complete pollinator, edaphic and distribution data are available showed that for sister species with overlapping distribution ranges, pollination system shifts are significantly associated with edaphic shifts. In contrast, there is no significant association between pollination system shifts and edaphic shifts for allopatric sister species. These results are interpreted as evidence for reinforcement.

ZUSAMMENFASSUNG

In den ersten drei Kapiteln dieser Dissertation präsentiere ich die Resultate meiner Arbeit über die Makroevolution. Bei den drei Hauptthemen handelt es sich um die Untersuchung der Kongruenz von Taxonomie und Verwandtschaftsverhältnissen, um die Rekonstruktion des Art-Stammbaumes, und um die Frage, welche Faktoren die Diversifizierung der Arten begünstigt haben.

Die Infragenerische Klassifikation

Im ersten Kapitel testete ich die Monophylie dreier Untergattungen von *Satyrium*. Obwohl die Monophylie von *Satyrium* durch morphologische und molekulare Daten gestützt wird, ist die Klassifikation innerhalb der Gattung wegen der grossen morphologischen Vielfalt sehr problematisch. Aus diesem Grund gibt es verschiedene Klassifizierungen. Die einzige infragenerische Klassifikation ist basiert auf einer Verwandtschaftsanalyse mit morphologischen Eigenschaften und erkennt die drei Untergattungen *Brachysaccium*, *Bifidum* and *Satyrium* an. Um die Monophylie dieser Untergattungen und den Status der problematischen Arten *S. rhynchanthum* and *S. pumilum* zu testen, wurden DNA-Sequenzen aus dem nukleären Genom (ITS1, 5.8S und ITS2) und aus dem Chloroplasten-Genom (*trnL* Intron und *trnL-F* Intergenic Spacer sowie Teile von *matK* und *trnK* Intron) verwendet und es wurde eine kombinierte Verwandtschaftsanalyse durchgeführt. Diese erfolgte unter Ausschluss von zwei Arten, die inkongruente Chloroplasten- und nukleäre Topologien besassen. *S. rhynchanthum* und *S. pumilum* wurden beide der Gattung *Satyrium* zugeordnet. Die *Parametric bootstrap*-Methode und *Bayesian posterior probabilities* zeigten, dass die drei untersuchten Untergattungen nicht monophyletisch sind. Stattdessen müssen andere Gruppierungen in Erwägung gezogen werden. Durch die Übertragen einiger morphologischen Merkmale auf den molekularen Stammbaum hat es sich gezeigt, dass es kein einziges Merkmal gibt, welches bei allen Mitgliedern eines *Clades* vorkommt oder für die Identifizierung eines *Clades* nützlich sein könnte.

Der Art-Stammbaum

Im zweiten Kapitel beschreibe ich in einem detaillierten Protokoll, wie man auf der Suche nach dem Art-Stammbaum mit phylogenetischer Inkongruenz umgehen kann. Das Verfahren gliedert sich in drei Schritte: 1. Identifizierung der Inkongruenz, 2. Aufspüren der

Gründe für die Inkongruenz, 3. Rekonstruktion des Art-Stammbaums. Zwei unterschiedliche phylogenetische Analysen von einem chloroplastidiären- und einem nukleären Datensatz mit 63 Arten von *Satyrium* brachten viele Inkongruenzen hervor. Ein *Incongruence length difference*-Test zeigte, dass viele davon nicht signifikant waren. In den übrigen signifikanten Fällen zeigten *Jack-knife* und *parametric bootstrap*-Experimente, dass nichtbiologische Artefakte wie unzureichendes *Taxon sampling* und *Long-branch attraction* als Gründe für die beobachtete Inkongruenz ausgeschlossen werden können. Um mögliche biologische Ursachen wie z.B. eine Vermischung von orthologen und paralogen Sequenzen, *Lineage sorting* oder Hybridisierung zu überprüfen, wurde die Anzahl Ereignisse gezählt, die im Nachhinein nötig wären, um das beobachtete Muster zu erzeugen. In den meisten Fällen, bei denen die Inkongruenz signifikant war, führte dies zu einer nahezu gleichen Anzahl von Ereignissen für alle verschiedenen potentiellen Ursachen. Nur für die drei südostasiatischen Arten, die monophyletisch sind, wurde Hybridisierung als mögliches Szenario bevorzugt. Diese Schlussfolgerung basiert auf der grossen Anzahl von Ereignissen, die angenommen werden müssten, um die alternativen Szenarien als Grund für die beobachtete Inkongruenz in Betracht zu ziehen. Die These eines hybriden Ursprungs dieses *Clades* wird auch durch morphologische Belege unterstützt. Der schliesslich rekonstruierte Art-Stammbaum ist ein Produkt einer kombinierten Analyse von Chloroplasten- und nukleären DNA-Sequenzen aller kongruenter Arten und nachträgliches Anhängen der inkongruenten *Clades* oder Akzessionen an den Stammbaum. Dieser Stammbaum stellt die derzeit beste phylogenetische Hypothese dar, und dient als Vorlage für weitere evolutionäre Untersuchungen.

Wechsel von Bestäubergruppe in der Kapflora

Im Kapitel 4 habe ich mich auf die reproduktive Isolation in der Kapflora konzentriert. Reproduktive Isolation kann entweder als Nebenprodukt von *divergent selection* oder durch *reinforcement* entstehen. Wir haben die Kapflora Südafrikas, die für ihre hohe Zahl von Bestäuberspezialisierung bekannt ist, als Modelsystem verwendet, um die potentielle Rolle von bestäubervermittelter Selektion im Artbildungsprozess zu testen. Vergleichende Analysen von 41 *sister-species pairs* (aus Geraniaceae, Iridaceae und Orchidaceae) für die komplette Daten zu Bestäubern, edaphischer Faktoren und Verbreitung vorhanden sind, haben gezeigt, dass für *sister-species pairs* mit überlappenden Verbreitungsgebieten, ein Wechsel der Bestäubergruppen signifikant mit dem Wechsel edaphischer Faktoren einherging. Im Gegensatz dazu gibt es keinen signifikanten Zusammenhang zwischen einer Verschiebung

von Bestäubergruppen und edaphischer Faktoren in allopatrischen *sister-species pairs*. Diese Ergebnisse werden als Beweiss für *reinforcement* interpretiert.

CURICULUM VITAE

Personal

Surname:	VAN DER NIET
First name:	Timotheüs
Date of birth:	23-04-1978
Nationality:	Dutch

Education

Highschool:	Gymnasium (Driestar College, Gouda), 1996
University:	M.Sc. in Biology (Wageningen University, Wageningen), 2000
Diploma subject:	Ecology
Title of Diploma thesis:	1. "The effect of hunting on the population dynamics of the red fox (*Vulpes vulpes*)" 2. "The effect of cattle grazing on breeding birds in a coastal dune system in the Netherlands."
Ph.D. thesis:	"Systematics and Radiation Patterns of *Satyrium* Sw. (Orchidaceae)" 2002-2006, supervised by Prof. H. P. Linder, University of Zurich
Grants:	SNF grant for one year post-doctoral fellowship at Wageningen University and Research Centre with Dr. F. Bakker, 2006-2007